

Lectins as Drug Targets: Functional, Structural,
and Pharmacological Insights into E-selectin and FimH

Inauguraldissertation

zur

Erlangung der Würde eines Doktors der Philosophie

vorgelegt der

Philosophisch-Naturwissenschaftlichen Fakultät

der Universität Basel

von

Roland Christopher Preston

aus Grossbritannien und Kemmental, TG

Basel, 2014

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Beat Ernst

Institut für Molekulare Pharmazie

Universität Basel, Klingelbergstrasse 50/70, CH-4055 Basel

Prof. Dr. Rudolf Glockshuber

Institut für Molekularbiologie und Biophysik

ETH-Hönggerberg, HPK E 17, Schafmattstrasse 20, CH-8093 Zürich

Basel, den 11. Dezember 2012

Prof. Dr. Jörg Schibler

Dekan

Acknowledgements

A doctoral thesis is rarely the work of a single individual. My work required the combined effort of a great team, reliable friends, and a supporting family and partner. I would like to take this chance to put the spotlight on all the people that have supported me throughout my PhD thesis and beyond.

First and foremost I would like to express my sincerest gratitude to Prof. Dr. Beat Ernst for not only giving me the chance to prove myself as a scientist, but also for giving me confidence in all situations. I truly enjoyed performing my masters and doctoral thesis in your group, almost 5 years of my life upon which I will look back on with a smile. I wish you great success in all of your future projects.

I further would like to thank to Prof. Dr. Rudolf Glockshuber for accepting to be the co-referee of my thesis.

During my time at the IMP I have seen many colleagues join and leave. I was welcomed into a friendly working environment by Gabi, Steffi, Jonas, Céline, Brian, Matthias, and Morena. I then had the pleasure to work with many of the IMP members on several projects. In the E-selectin project, I would like to express my gratitude to Dr. Said Rabbani for introducing me to the world of E-selectin and molecular biology during my Masters thesis, Dr. Katrin Lemme, Dr. Florian Binder, Mirko Zierke, Dr. Martin Smieško, and Bea Wagner. In the FimH project I thank Dr. Said Rabbani, Adam Zalewski, Dr. Meike Scharenberg, Deniz Eris, and Daniela Abgottspon. In the PapG project I would like to thank Dr. Katja Stangier and Guilo Navarra. Further huge thanks go also to all the rest of the IMP for making it a great place to work at: Jacqueline, Simon, Lijuan, Fan, Olly, Xiaohua, Claudia, Sameh, Kathi, Wojtek, Pascal, and Anja. I wish you all best of luck at the IMP.

A one-of-a-kind friendship evolved between me and Meike since day one. Thank you so very, very much for being there when things did not go as planned, when they went as planned, when I needed motivation, when I just needed to talk, when I required your scientific advice and and and... Deniz Eris also deserves special mentioning. Many thanks to you for all the tech-talk and the daily fun. It is unfortunate that the members of our little lab-family have to go separate ways, but we will always remain friends.

I further would like to point out the help I received from my Master students Silvan Röthlisberger, Deniz Eris, and Fabian Hulliger. I was fortunate to have you all fast learning, reliable, hard working, and always fun students.

Acknowledgements

Many thanks go also to our external partners: Prof. Dr. Viola Vogel together with Dr. Jens Möller and Philippe Emge from the ETH Zürich for introducing me to the flow chamber experiments. Prof. Dr. Gerhard Klebe and his team for giving me first insights into protein crystallography. Dr. Andreas Menzel from the Swiss Light Source for aiding in SAXS experiments.

I would like to stress the immense help I received from Prof. Dr. Timm Maier and Dr. Roman Jakob from the Biocenter (University of Basel) during the various crystallization projects. I will never forget the many hours you have spent to introduce me to the fascinating world of protein-crystals out of sheer goodwill.

I owe sincere thankfulness to my parents for always believing in me and for supporting me in all imaginable ways over the past 28 years. My warmest thanks go also to my sister, my brother, my brother-in-law, and my two nephews (who will be scientists oneday).

Finally, I would like to thank Balbina for sensing when I needed your care during the ups-and downs during my doctorate, and simply for the amazing time we have together.

Abstract

Lectins are carbohydrate-binding proteins found throughout nature in plants, viruses, pro- and eukaryotes. Their carbohydrate specificity as well as their biological functions are highly diverse. Lectins in humans primarily serve as cell-surface receptors and are mainly involved in the immunogenic processes. Bacterial lectins are often involved in conferring pathogenesis by binding to carbohydrates of the host. In this thesis, structural, functional, and pharmacological insights on the mammalian lectin **E-selectin** and the bacterial lectin **FimH** are given.

E-selectin is a C-type lectin involved in leukocyte recruitment during inflammation by binding to the tetrasaccharide ligand sialyl Lewis^x (sLe^x). It is involved in numerous diseases, e.g. asthma, psoriasis, stroke, rheumatoid arthritis or cancer metastasis. Targeting E-selectin with glycomimetic antagonists is therefore in the focus of drug discovery. The following aspects of E-selectin were investigated:

- **Publication 1** – The thermodynamic driving forces of the interaction of E-selectin with sLe^x and glycomimetics thereof were investigated. We demonstrated that sLe^x binding is an entropy-driven process, which is an uncommon feature of carbohydrate-lectin interactions.
- **Manuscript 1** – The co-crystallization of E-selectin with sLe^x or a more potent glycomimetics thereof revealed a previously unseen induced fit of the binding site involving alterations in the first two domains. We showed that this induced fit occurs in solution and discuss the physiological relevance.
- **Manuscript 2** – A flexible and a pre-organized E-selectin antagonist were characterized for their kinetic and thermodynamic properties, which revealed an unexpected loss in entropy for the pre-organized antagonist. Co-crystallization with a series of antagonists revealed the reason for this behavior.
- **Publication 2** – The single nucleotide polymorphism which leads to the S128R mutation in E-selectin has been correlated with an increased risk of developing various diseases. We investigated the binding behavior of this mutant and demonstrated that glycomimetics are efficacious in inhibiting E-selectin-S128R mediated binding.

- **Manuscript 3** – Mice are able to express the *N*-glycolyl form of sLe^x, unlike humans. Therefore, the specificity of murine E-selectin might be altered. We investigated the binding specificity of murine E-selectin and evaluated the potency of antagonists designed for human E-selectin. We confirmed the efficacy of E-selectin antagonists towards murine E-selectin, thus demonstrating the validity of mouse models.

The bacterial lectin **FimH** is presented by uropathogenic *E. coli* (UPEC) on the tip of type 1 pili and mediates the adhesion to mannosylated structure in the lower urinary tract. This interaction allows UPEC to colonize the bladder, the initial step in bladder infection. Mannoside-based FimH antagonists are under investigation as treatment for bladder infections.

- **Manuscript 4** – The goal of a drug discovery program aimed to develop a treatment for urinary tract infections is to identify high-affinity, orally available, and safe FimH antagonists. Starting from the carboxylate substituted biphenyl α -D-mannopyranoside, affinity as well as the relevant pharmacokinetic parameters (solubility, permeability, renal excretion) could be substantially improved by a bioisosteric approach.
- **Manuscript 5** – To comprehend and further develop potent FimH antagonists, structural data on ligand-protein interaction is essential. In this manuscript we present the X-ray co-crystal structures of FimH with three antagonist classes for which structural data were unavailable to date and provide an explanation for the observed entropy-enthalpy compensation by NMR.
- **Manuscript 6** – Crystallographic studies of FimH with alkyl- or aryl-substituted α -D-mannopyranosides have demonstrated alternative binding poses with differing involvement of the residues Tyr48 and Tyr137 at the binding site entrance. Thermodynamic and molecular modeling analysis provided insights into the importance of the tyrosine-gate.
- **Manuscript 7** – Several mutations of FimH are found in clinical isolates, which influence the binding phenotype of FimH by altering the interaction of the two FimH domains (lectin- and pilin domain). To date, FimH antagonists have never been tested on clinically relevant FimH variants. We demonstrated that antagonist affinity correlated with the binding behavior of different FimH variants.

Table of Contents

1. Introduction to the Lectin Family	1
2. The Lectin E-Selectin	5
2.1. Introduction to the Selectins	5
2.2. Publication 1. Sialyl Lewis ^x : A “Pre-Organized Water Oligomer?”	29
2.3. Manuscript 1. E-Selectin Ligand Complexes Adopt an Extended High-Affinity Conformation	51
2.4. Manuscript 2. Acid Pre-Organization of a Glycomimetic E-selectin Antagonist	79
2.5. Publication 2. Implications of the E-selectin S128R Mutation for Drug Discovery	101
2.6. Manuscript 3. <i>In vitro</i> Comparison of Murine and Human E-selectin Binding to Natural Ligands and Antagonists	117
3. The Lectin FimH	129
3.1. Introduction to FimH	129
3.2. Manuscript 4. FimH Antagonists - Bioisosteres to Improve the <i>in vitro</i> and <i>in vivo</i> PK/PD Profile	149
3.3. Manuscript 5. The tyrosine gate of FimH – conformational analysis by NMR and X-ray	219
3.4. Manuscript 6. Structural and Thermodynamic Characterization of the Interaction between the Bacterial Adhesin FimH and Mannosides: The Importance of the Tyrosine Gate	247
3.5. Manuscript 7. Naturally Occurring Variations of the FimH Adhesin: Impact on Antagonist Binding	281

Chapter 1. – The Lectin Family

1. Introduction to the Lectin Family

The lectins are a family of proteins that share specific binding towards carbohydrates. They are widespread throughout nature, as they exist in plants (*e.g.* concanavalin A first isolated from *Canavalia ensiformis*¹), viruses (*e.g.* hemagglutinin²), and in pro- and eukaryotic cells^{3,4}. Since their initial discovery in 1888, numerous lectins with a wide range of functions have been identified. While initially plant lectins have been thoroughly investigated, the main focus has shifted towards animal lectins since the 1970s, in particular human lectins.

1.1. Human Lectins

Lectins in humans cover a variety of functions, *e.g.* in cell-cell adhesion, cell signaling, host immune response, embryonic development *etc*^{5,6}. Based on their sequence homology and their functional relationship, mammalian lectins have been classified into four major categories.

- **C-type lectins** which share a common requirement of one or more Ca^{2+} ions for carbohydrate recognition and binding, while having a diverse carbohydrate specificity⁷. The family of selectins, the Dendritic Cell-Specific Intercellular adhesion molecule-3-Grabbing Non-integrin (DC-SIGN), and asialoglycoprotein receptors are among the most thoroughly investigated C-type lectins.
- **Galectins** (formerly known as S-type lectins) with specificity towards β -galactosides share conserved amino acid residues among their carbohydrate binding domains, which are directly involved in carbohydrate binding⁸. They have widespread functions, since they are involved in immune response, inflammation, apoptosis and in embryonic development.
- **P-type lectins** which require mannose-6-phosphate (M6P) for binding affinity⁹. Two P-type lectins have been identified: The cation dependent receptor (CD-MPR), and the cation-independent receptor (CI-M6P). Both are involved in the generation of functional lysosomes.
- **I-type lectins** that belong to the immunoglobulin superfamily and contain at least one immunoglobulin-like fold¹⁰. The family of Siglecs is the most prominent group of I-type lectins, which all share a requirement of sialic acid (NeuNAc). Their function ranges from cell-cell adhesion in the host immune response (Sialoadhesin) to the inhibition of myelination in neuronal growth (Myelin-associated glycoprotein).

As summarized above, mammalian lectins are present in a wide variety, each with distinct function and carbohydrate specificities and affinities. Dysfunction or override of some of these lectins can lead to diseases, *e.g.* B-cell non-hodgkin lymphoma, where Siglec-2 is involved¹¹ or DC-SIGN, which is exploited as an entry point by numerous microbials, such as the HI-virus¹². Furthermore, the selectin family is also involved in numerous diseases. A more detailed description on their function in health and disease is noted in Section 2.1.

1.2. Microbial Lectins

Microbial lectins are mainly located on bacterial fimbriae and mediate the initial adhesion to glycoproteins, glycolipids or proteoglycans presented on endo- and epithelial cells of the host¹³. Therefore, microbial lectins are commonly referred to as “adhesins” and represent important virulence factors for bacteria. The best characterized adhesins of *Escherichia coli* (*E. coli*) to date are:

- **FimH**, located on type 1 pili with a specificity towards mannose structures presented on the urothelium¹⁴,
- **PapG**, located on P pili with a specificity towards galabiose (Gal α 1–4Gal β) presented in the upper urinary tract¹⁵,
- **F17G**, located on F17 pili with a specificity towards *N*-acetylglucosamine (GlcNAc) presented on intestinal mucosal cells¹⁶.

The interaction of these adhesins with specific carbohydrates allows pathogenic *E. coli* to adhere and subsequently to invade host cells which present these carbohydrates on their surfaces. For example, *E. coli* initially adhere to the lower urinary tract via FimH, and, at a later stage of infection, to the upper urinary tract using PapG. A detailed overview on type 1 pili and FimH is given in Section 3.1.

1.3. References for Chapter 1.

1. Sumner, J.B., Gralen, N. & Quensel, I.B.E. The molecular weights of urease, canavalin, concanavalin A and concanavalin B. *Science* **87**, 395-396 (1938).
2. Weis, W., Brown, J.H., Cusack, S., Paulson, J.C., Skehel, J.J. & Wiley, D.C. Structure of the influenza virus haemagglutinin complexed with its receptor, sialic acid. *Nature* **333**, 426-31 (1988).
3. Sharon, N. Bacterial Lectins, Cell-Cell Recognition and Infectious-Disease. *FEBS Lett.* **217**, 145-157 (1987).
4. Drickamer, K. & Taylor, M.E. Biology of Animal Lectins. *Annu. Rev. Cell Biol.* **9**, 237-264 (1993).
5. Crocker, P.R. & Feizi, T. Carbohydrate recognition systems: Functional triads in cell-cell interactions. *Curr. Opin. Struct. Biol.* **6**, 679-691 (1996).
6. Poirier, F. & Kimber, S. Cell surface carbohydrates and lectins in early development. *Mol. Hum. Reprod.* **3**, 907-18 (1997).
7. Zelensky, A.N. & Gready, J.E. The C-type lectin-like domain superfamily. *FEBS J.* **272**, 6179-217 (2005).
8. Barondes, S.H., Cooper, D.N., Gitt, M.A. & Leffler, H. Galectins. Structure and function of a large family of animal lectins. *J. Biol. Chem.* **269**, 20807-10 (1994).
9. Hancock, M.K., Haskins, D.J., Sun, G. & Dahms, N.M. Identification of residues essential for carbohydrate recognition by the insulin-like growth factor II/mannose 6-phosphate receptor. *J. Biol. Chem.* **277**, 11255-64 (2002).
10. Powell, L.D. & Varki, A. I-type lectins. *J. Biol. Chem.* **270**, 14243-6 (1995).
11. Tedder, T.F., Poe, J.C. & Haas, K.M. CD22: a multifunctional receptor that regulates B lymphocyte survival and signal transduction. *Adv. Immunol.* **88**, 1-50 (2005).
12. Geijtenbeek, T.B.H., Kwon, D.S., Torensma, R., van Vliet, S.J., van Duijnhoven, G.C.F., Middel, J., Cornelissen, I.L.M.H.A., Nottet, H.S.L.M., KewalRamani, V.N., Littman, D.R. et al. DC-SIGN, a dendritic cell-specific HIV-1-binding protein that enhances trans-infection of T cells. *Cell* **100**, 587-597 (2000).
13. Mouricout, M. Interactions between the enteric pathogen and the host. An assortment of bacterial lectins and a set of glycoconjugate receptors. *Adv. Exp. Med. Biol.* **412**, 109-23 (1997).
14. Jones, C.H., Pinkner, J.S., Roth, R., Heuser, J., Nicholes, A.V., Abraham, S.N. & Hultgren, S.J. FimH adhesin of type 1 pili is assembled into a fibrillar tip structure in the Enterobacteriaceae. *Proc. Natl. Acad. Sci. U. S. A.* **92**, 2081-5 (1995).
15. Kuehn, M.J., Heuser, J., Normark, S. & Hultgren, S.J. P pili in uropathogenic E. coli are composite fibres with distinct fibrillar adhesive tips. *Nature* **356**, 252-5 (1992).
16. Bertin, Y., Martin, C., Oswald, E. & Girardeau, J.P. Rapid and specific detection of F17-related pilin and adhesin genes in diarrheic and septicemic Escherichia coli strains by multiplex PCR. *J. Clin. Microbiol.* **34**, 2921-8 (1996).

Chapter 2. – The Lectin E-selectin

2.1. Introduction to the Selectins

The selectins are type-1-membrane adhesion molecules consisting of the three members E-, L-, and P-selectin^{1,2}. Since they have a functional requirement for a Ca^{2+} ion, the selectins belong to the class of C-type lectins. Their expression is restricted to bone-marrow derived cells and the vascular endothelium³. The main function of selectins is the guidance of leukocytes to specific tissues by allowing initial contacts (*i.e.* tethering) between free flowing leukocytes from the bloodstream and endothelial cells, *e.g.* during inflammation^{4,5} and homing of lymphocytes⁶.

- E-selectin (also termed CD62E, ELAM-1, LECAM-2) is uniquely expressed on postcapillary vascular endothelial cells upon an inflammatory stimulus. Its main role is the capturing of free flowing leukocytes at a site of inflammation.
- L-selectin (also termed CD62L, LAM-1, LECAM-1) is constitutively expressed on blood neutrophils, monocytes and on most blood-borne T- and B-cells⁷. Interaction of L-selectin with ligands on high endothelial cells (HEV) allows leukocytes to be captured, a prerequisite for their entry into secondary lymphoid organs. This process is known as lymphocyte homing⁸.
- P-selectin (also termed CD62P, LECAM-3, GMP-140, PADGEM) is displayed on platelets⁹ as well as on vascular endothelial cells¹⁰ upon induction. As for E-selectin, the display of P-selectin on endothelial cells initially captures leukocytes from the blood stream. Expression on platelets leads to platelet aggregation during blood clot formation¹¹.

2.1.1. Overall Structure of the Selectins

All three selectins share a similar overall protein structure as shown in Figure 2-1. The carbohydrate recognition domain (CRD), also referred to as the lectin domain, is located at the N-terminus, with the largest distance to the cell surface. This domain of 120–130 amino acids contains the ligand binding site and coordinates the Ca^{2+} ion¹².

The lectin domain is C-terminally connected to the epidermal growth factor-like (EGF-like) domain, a relatively small domain of approximately 36 amino acids. The function of this domain is less understood. The binding activity of the lectin domain however requires the presence of the EGF-like domain², since deletion of the EGF-like domain strongly impairs selectin affinity^{13,14}. Therefore, the two N-terminal domains are considered as the minimal requirement for selectin activity. As the EGF-like domain does not directly interact with the

carbohydrate ligand, it is proposed that it might stabilize the lectin domain. The X-ray structure of the E-selectin lectin- and EGF-like domain supports this notion, in which interactions of residues 135–139 of the EGF-like domain with residues of the lectin domain are observed (PDB code 1ESL)¹⁵.

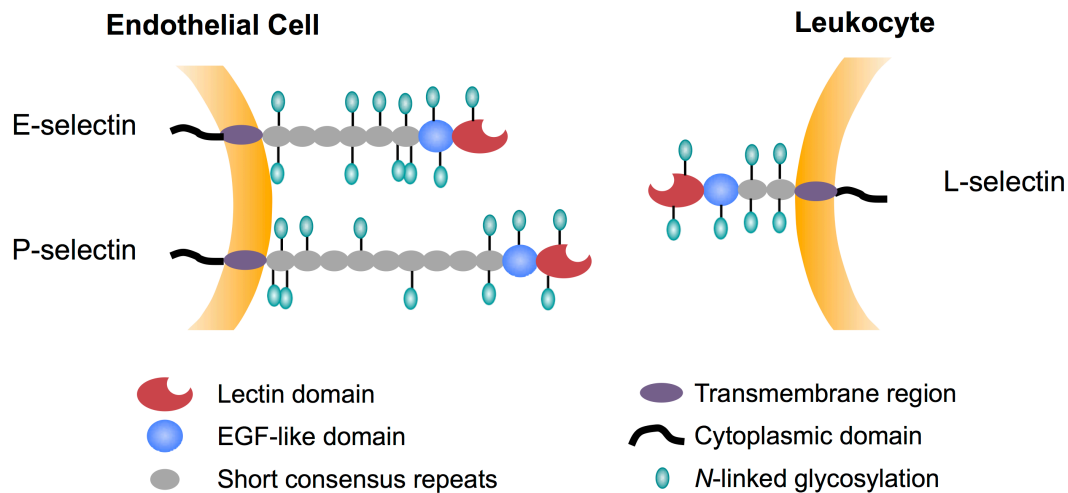


Figure 2-1. Overall structure of human selectins. The N-terminal carbohydrate recognition domain allows Ca^{2+} dependent binding. The lectin and EGF-like domain together are the minimal requirement for selectin binding activity. All selectins possess extensive *N*-linked glycosylation.

The EGF-like domain is followed by a variable number of short consensus repeats (SCR), also called sushi domains or complement repeat domains, each with ~60 amino acids and three disulfide bonds. In humans, P-selectin possesses nine, E-selectin six, and L-selectin two SCRs¹⁶. Their function remains unclear, since they are neither directly involved in ligand binding, nor in selectin oligomerization. Using cell-based assays it has been shown that increasing numbers of SCRs in E-selectin lead to augmented binding affinity¹⁷, and that substitution of SCRs among selectins alters their specificity¹⁸. The presence or absence of both SCRs in L-selectin has also been correlated with binding affinity¹⁹. Kolbinger *et al.*, however, did not confirm that the SCRs determine ligand specificity²⁰. A recent study suggests that SCR might act as springs under mechanical force²¹, as is present for leukocyte tethering under flow conditions. Finally, SCRs might merely act as spacer for the lectin domain, in order to protrude from the cell surface into the glycocalyx²².

The selectins are anchored to the cell membrane by a ~23 amino acids transmembrane region. The C-terminal cytoplasmic domain consists of 17–35 amino acids, is involved in signaling functions^{23,24}, and interacts with the cytoskeleton²⁵.

The primary sequence homology between human selectins is relatively low. For the lectin domain it is ~52%, for the EGF-like domain ~47%, and for the SCRs ~35%. The transmembrane and cytoplasmic domains do not share any homology, suggesting various effector functions of selectins.

2.1.2. Ligands of Selectins

While a variety of glycoproteins have been identified to serve as selectin ligands, they all share a common minimal binding epitope, namely sialyl Lewis^x (sLe^x) and its regioisomer sialyl Lewis^a (sLe^a) as shown in Figure 2-2. These two tetrasaccharides are recognized by all three selectins^{26,27}. The glycoproteins thereby act as scaffolds bearing glycans terminally conjugated with sLe^x. The affinity of sLe^x itself to selectins is very weak, with 0.3–1.1 mM for E-selectin, and 7–9 mM for P-selectin²⁸.

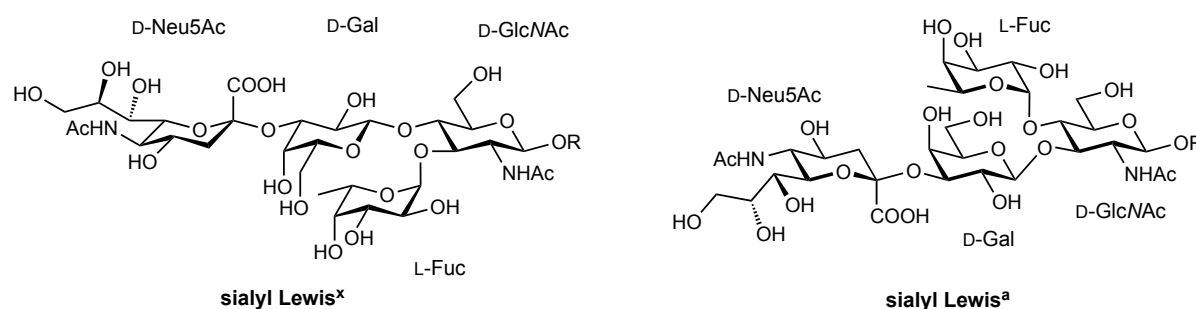


Figure 2-2. The tetrasaccharides sialyl Lewis^x (Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc) or sialyl Lewis^a (Neu5Ac(α2-3)Gal(β1-3)[Fuc(α1-4)]GlcNAc) are the minimal binding epitopes for all selectins.

Glycoproteins bearing sLe^x (Figure 2-3) in some cases also provide additional interactions with selectins, which may influence specificity. A well studied natural selectin ligand is P-selectin glycoprotein ligand-1 (PSGL-1), a homodimeric glycoprotein presented on the tips of the microfolds on leukocytes²⁹. While selectins in most cases bind with weak affinity to sLe^x bearing glycoproteins, a K_D value of 320 nM for the interaction of P-selectin with PSGL-1 was measured using surface plasmon resonance (SPR)³⁰. This high affinity is not only a result of the interaction of sLe^x with P-selectin, but with additional interactions of sulfated tyrosine residues on PSGL-1. The lack of secondary binding sites for sulfated groups leads to a vastly reduced binding affinity of E-selectin with PSGL-1^{26,31}, which nonetheless is considered as an E-selectin ligand. Another well investigated glycoprotein ligand is E-selectin ligand-1 (ESL-1), a 150 kDa, non-sulfated glycoprotein located on leukocytes that binds exclusively to E-selectin with moderate affinity 62 μM³²⁻³⁴.

L-selectin has also exclusive sLe^x bearing glycoprotein ligands, *e.g.* addressin (MAdCAM-1)^{35,36}, CD34³⁷, endomucin³⁸, endoglycan³⁹, and glycosylation-dependent cell adhesion molecule-1 (GlyCAM-1)⁴⁰. Most of these L-selectin ligands are sialomucins and are expressed on endothelial cells, enabling leukocyte trafficking.

By regulating the sLe^x biosynthesis through glycosyltransferases and the modulation of protein sulfation by sulfo-transferases, leukocytes can effectively control their selectin specificity, and hence direct their adherence to various tissues³.

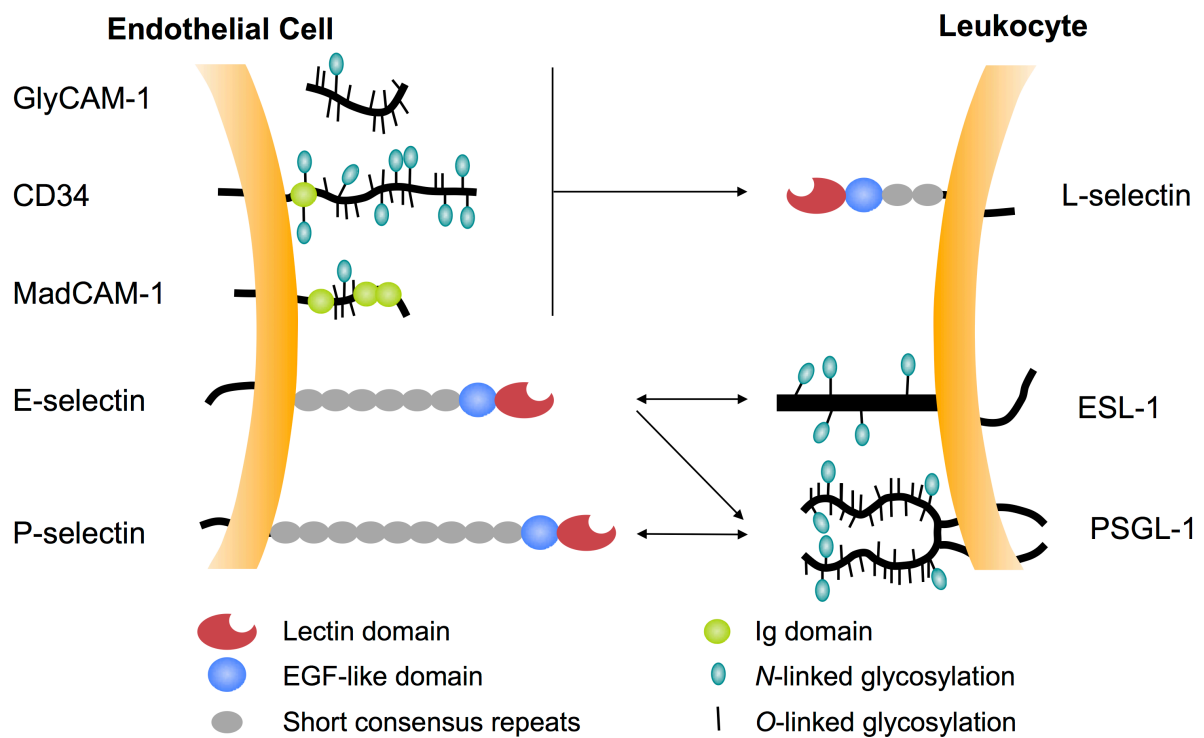


Figure 2-3. Overview of selectin glycoprotein ligands all terminally bearing sLe^x. The homodimeric PSGL-1 is a ligand for both, E- and P-selectin. ESL-1 binds exclusively to E-selectin while L-selectin has its own set of ligands.

2.1.3. The Inflammatory Cascade

Leukocytes are cells of the immune system circulating in the blood stream that defend against infections and foreign materials. Their mechanism of action requires leukocytes to traffic to the site of tissue damage. This is achieved by a well defined sequence of events termed the inflammatory cascade (Figure 2-4).

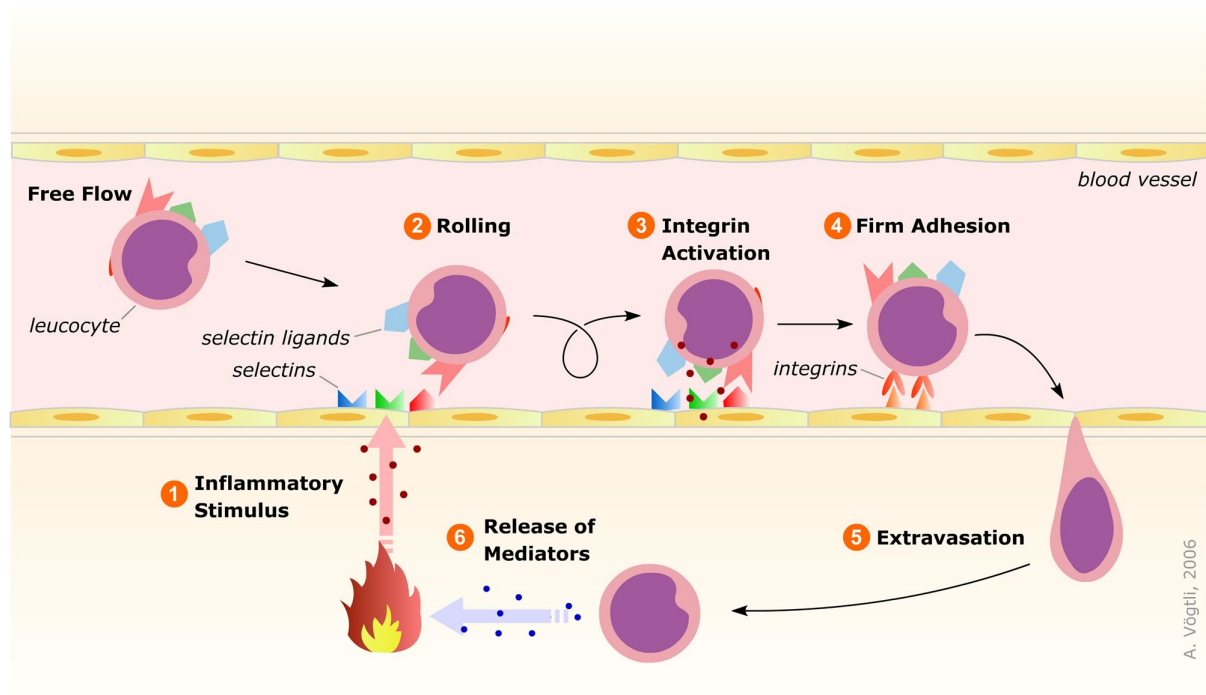


Figure 2-4. Schematic representation of the inflammatory cascade (courtesy of Alexander Vögtli). An inflammatory stimulus (1) is caused by the release of cytokines upon tissue damage or irritation. This triggers the expression of selectins on the vascular endothelium (2), upon which leukocytes are able to initially tether and roll along the surface. Activation of integrins (3), allowing further cell-cell contacts, translates the rolling motion into a firm adhesion, after which leukocytes extravasate (5) into the tissue where they resolve the cause of inflammation.

In the first step, pro-inflammatory cytokines are released at the site of tissue lesion. This initiates the presentation of E- and P-selectin on vascular endothelial cells. P-selectin is rapidly displayed on the cell surface, since it is stored in Weibel-Palade bodies (granules) within the endothelial cells. Upon the release of thrombin, histamine, activators of protein kinase C or complement fragments, these granules fuse with the cell membrane. Full display of P-selectin is usually complete after 10–20 minutes^{41,42}. Furthermore, P-selectin is regulated on the level of mRNA transcription by lipopolysaccharides (LPS), tumor necrosis factor- α (TNF- α), and various interleukins, which ensures stable display after several hours at a later stage of inflammation⁴³⁻⁴⁵.

E-selectin, is solely regulated on the transcription level by the NF κ B-pathway and is expressed upon increased levels of LPS, TNF- α , and interleukin-1 β (IL-1 β)^{46,47}. Therefore, E-selectin is displayed only at later stages of inflammation, usually within 4–6 hours after stimulation. The late engagement of E-selectin was confirmed in E-selectin knockout mice, where leukocyte tethering in comparison to the control group was impaired only several hours after stimulation⁴⁸.

L-selectin, unlike E- and P-selectin being located on leukocytes, also plays a role in leukocyte tethering in the early stages of inflammation by interacting with glycoproteins presented on endothelial cells⁵. In addition, L-selectin interacts with already tethered leukocytes, further increasing the number of adherent cells by this secondary tethering⁴⁹.

Once leukocytes are initially captured they roll along the vasculature. This rolling motion is regulated by the force the blood flow exerts on the leukocyte, and by the kinetics of selectin-ligand interaction (see Section 2.1.5.). While L-selectin allows high-velocity rolling, P-selectin shows intermediate, and E-selectin only slow rolling. It is thought that the slow rolling on E-selectin facilitates the transition into firm adhesion⁴.

Selectins alone however are not sufficient to enable leukocyte to transmigration into the surrounding tissue. The display of integrins on leukocytes is triggered via G-protein coupled activation once in contact with the endothelial cells. The interaction of integrins, *e.g.* the lymphocyte function-associated antigen-1 (LFA-1) for lymphocytes or macrophage-1 antigen (Mac-1) for macrophages, with the intercellular adhesion molecule-1 (ICAM-1) on endothelial cells mediates the complete arrest of leukocytes⁴. There is a co-dependence between the integrin system with the selectin system, since knockout animals of either system have impaired leukocyte extravasion⁵⁰.

In order to resolve the cause of inflammation, the immobilized leukocytes must overcome the endothelial cell layer and penetrate into the tissue. This is achieved with the involvement of numerous factors, which requires cytoskeletal rearrangement in both, the leukocytes and the endothelial cells. Two mechanisms have been identified, a paracellular pathway, where leukocytes transmigrate between cells, and a transcellular pathway⁵¹, in which leukocytes pass through endothelial cells. Both pathways are thought to occur during transmigration. Once in the tissue, chemokine gradients (IL-8) attract leukocytes towards the site of infection⁴.

2.1.4. Selectins in Pathology

Selectins are responsible for trafficking of leukocytes in a well regulated fashion. However, dysregulation of selectins can lead to various states of disease. This often involves exaggerated leukocyte recruitment without an inflammatory stimulus, leading to destruction of healthy tissue⁵². This occurs in a variety of cardiovascular diseases such as atherosclerosis⁵³, ischemia (*e.g.* myocardial infarction⁵⁴), reperfusion injury⁵⁵⁻⁵⁷, hypertension⁵⁸, and diabetes⁵⁹. Also immune diseases are associated with selectins, *e.g.* rheumatoid arthritis⁶⁰, systemic lupus erythematosus⁶¹, and asthma bronchiale⁶². Furthermore, selectin contributions have been found to be relevant in cancer metastasis and in sickle cell disease⁶³.

Single nucleotide point mutations (SNP) of E-selectin have also been associated with diseases. The S128R mutation, which is located in the EGF-like domain, has been correlated with an increased risk and severity of a number of diseases, *e.g.* myocardial infarction⁶⁴, severe asthma⁶⁵, colon metastasis formation⁶⁶ *etc.* Another SNP in E-selectin which causes the Leu554Phe mutation in the transmembrane region is associated with atherosclerosis⁶⁷ and hypertension⁶⁸.

While selectins themselves are usually not the causative agent of such diseases, their function as adhesion molecules sustains or even aggravates certain pathogenic processes. Therefore, inhibition of selectin function is considered as a highly relevant drug target.

2.1.5. Selectin Binding Properties

Kinetics

The function of leukocytes is to allow the tethering of fast flowing leukocytes from the blood stream and to enable rolling along the vasculature. This requires specific binding properties regarding kinetics of selectins with their respective ligands (see Figure 2-5). A rapid and specific interaction is required during initial tethering and rolling (described by the association constant k_{on}) at the leading edge of the leukocyte. The bond lifetime, on the other hand, must be short in order to break bonds at the trailing edge of the leukocyte and allow rolling (described by the dissociation constant k_{off}). A balance of k_{on} and k_{off} is therefore of great importance for proper selectin function⁶⁹. In addition, the surface densities of selectins and their ligands was also shown to influence rolling behavior^{70,71}.

The K_D value describes the equilibrium of binding in a static environment, which is never reached under flow conditions^{72,73}. The mean velocity of rolling leukocytes is 10–100 fold lower than that of the blood flow. Surface plasmon resonance (SPR) studies were performed to investigate the kinetics of selectin binding. For the E-selectin–ESL-1 interaction, a K_D of 62 μM was determined with a k_{off} rate of 4.6 s^{-1} and a k_{on} rate of $7.4 \cdot 10^4 \text{ M}^{-1} \text{ s}^{-1}$. This translates into a half-life of only 0.15 seconds. For the P-selectin–PSGL-1 interaction k_{on} of $4.4 \cdot 10^6 \text{ M}^{-1} \text{ s}^{-1}$ and a k_{off} of was 1.4 s^{-1} ($K_D = 0.3 \text{ }\mu\text{M}$). The half-life of 0.5 seconds is in the same order of magnitude as for the E-selectin–ESL-1 interaction, albeit the additional interactions of sulfated tyrosines PSGL-1. The faster k_{on} of P-selectin suggests differential role of E- and P-selectin: while P-selectin is thought to mediate the very first leukocyte interaction, E-selectin is proposed to mediate the transition of fast to slow rolling and finally the full arrest of the cell³⁴. E-selectin itself has only low potential in capturing leukocytes from the bulk flow⁷⁴.

Selectin Binding under Mechanical Force

Selectin–ligand binding has been shown to be modulated under tensile force (see Figure 2-5). This force arises by the drag of leukocytes once bound to the endothelial cells, which pulls along the selectin–ligand axis. It has been reported that optimal selectin-mediated adherence and rolling requires a minimal shear threshold. The physiological wall shear stress in postcapillary venules is 1–10 dynes/cm^2 . For L-, E-, and P-selectin, the minimal wall shear stress threshold is 0.8, 0.25, and 0.5 dynes/cm^2 , respectively⁷⁵. This is explained by stabilization of selectin–ligand interaction under tensile force. While the interaction is unstable under static conditions, stability is enhanced up to a certain point, after which stability decreases. This phenomenon is termed catch-bonds, in contrast to conventional slip bonds where bond lifetimes decrease under force. This binding behavior of the selectins has been studied using atomic force microscopy⁷⁶, biomembrane force probes⁷⁷, and flow chamber assays⁷⁸.

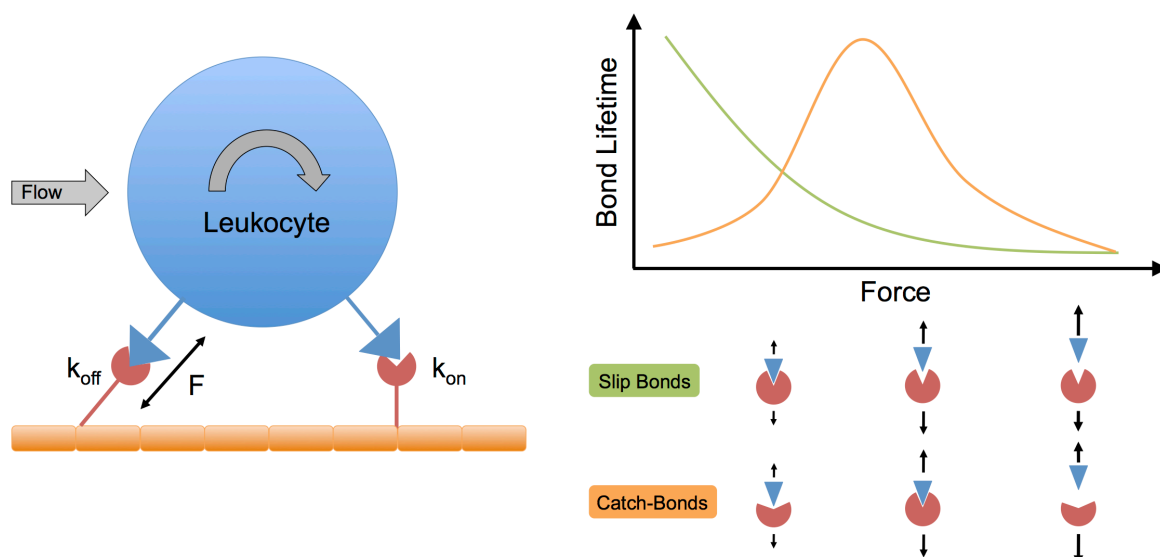


Figure 2-5. Binding characteristics of selectins. Left: To enable rapid capturing of leukocytes, selectins must provide fast k_{on} kinetics at the leading edge of the leukocytes. The breaking of bonds at the trailing edge is governed by the dissociation rate (k_{off}) and allows for the rolling motion. Right: Selectins display catch-bond behavior. Most protein-ligand interactions are weakened when tensile force is applied to the bond (slip-bonds, green). For catch-bonds, bond lifetimes are initially increased to a certain threshold (orange). Catch-bonds are accompanied by a change in binding site conformations.

Two models were developed to explain the catch-bond phenomenon for selectins. i) Based on two distinct crystal structures of P-selectin (see Section 2.1.7.), it is thought that allosteric alterations of selectins might strengthen the interaction by forcing P-selectin from a bent to an extended conformation and altering the binding site in a way that increases binding affinity^{79,80}. Indeed, mutational studies which force selectins into the extended conformation under static conditions have shown to increase binding affinities^{81,82}. ii) In the sliding-rebinding model, on the other hand, it is thought that alignment of the binding interface along the applied force allows the ligand to slide along the surface and continuously form new interactions⁸³.

2.1.6. Molecular Insights of the E-selectin–sialyl Lewis^x Interaction

The interaction of sLe^x with E- and P-selectin has been characterized with numerous methods. Prior to the elucidation of E- and P-selectin X-ray structures with sLe^x, the hydroxyl groups of L-fucose have been identified as important pharmacophores^{84,85}. Furthermore, the 4- and 6-hydroxyl groups of D-galactose have been identified as crucial for binding⁸⁶⁻⁸⁸, as well as the carboxylic acid of D-Neu5Ac⁸⁵. The D-GlcNAc moiety has been shown not to directly interact with the protein⁸⁹⁻⁹¹, which rather acts as spacer for the proper

orientation of pharmacophores⁹². Transferred Nuclear Overhauser Enhancement-NMR (tr-NOE-NMR) confirmed all previous studies on the sLe^x pharmacophores in the protein bound conformation^{28,93-95}. Mutational studies on E-selectin revealed the relevant amino acid residues. Erbe *et al.* propose that the residues Ser43, Tyr48, Arg97, Lys111, and Lys113 are crucial residues in sLe^x binding¹².

The crystal structures of E- and P-selectin published in 2000 by Somers *et al.* confirmed the previously made assumptions on sLe^x pharmacophores and the residues involved in binding⁷⁹. The following main observations for the E-selectin structure soaked with sLe^x (PDB code 1G1T)⁷⁹ are made (Figure 3-6):

- The L-fucose moiety coordinates with its 3- and 4-hydroxyl groups to the Ca²⁺ ion, which is also coordinated by residues of the protein. The 2-hydroxyl group of L-fucose mediates a water-bridged interaction to the protein.
- The GlcNAc moiety shows no interaction to the protein.
- The carboxylate group of Neu5Ac provides a salt bridge to Arg97 and Tyr48.
- The oxygen between the glycosidic bond of Neu5Ac and D-galactose interacts with Arg97 side chain.
- D-galactose provides hydrogen bonds with its 4- and 6-hydroxyl groups.
- The binding surface has a large surface area of 275 Å².

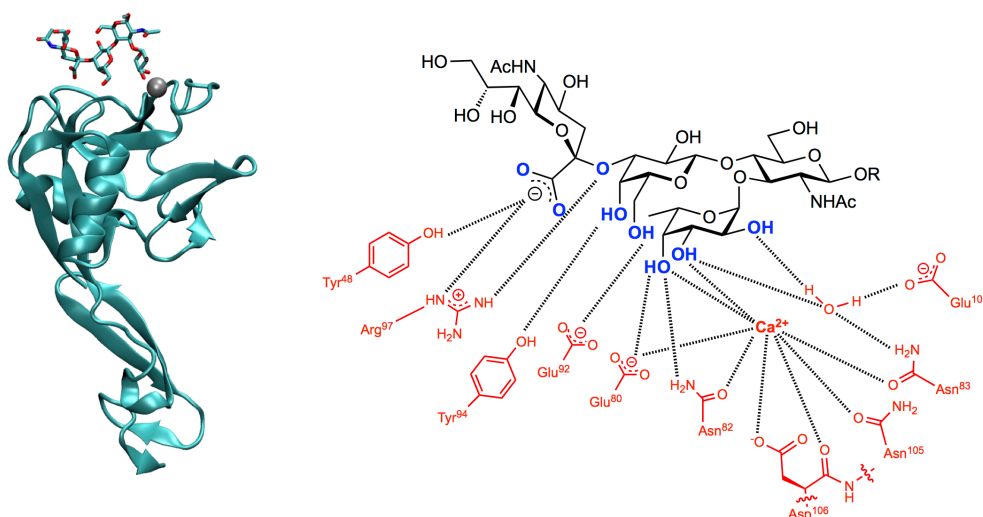


Figure 2-6. Left: sLe^x in the binding site of E-selectin (PDB code 1G1T)⁷⁹. The Ca²⁺ ion is depicted as grey sphere. Right: Schematic representation of the E-selectin–sLe^x interaction as observed in the soaked crystal structure. Relevant pharmacophores are marked in blue, while amino acid residues are marked in red.

Superimposing the structures of E-selectin or P-selectin (PDB code 1G1Q)⁷⁹ soaked with sLe^x reveals near identical protein conformation. Furthermore, the bioactive conformation of sLe^x in both structures is identical and is one of two conformations found in solution⁹⁶⁻⁹⁸. sLe^x is therefore a rigid molecule in solution and requires only slight conformational changes upon binding, an unusual feature for carbohydrates⁹⁹.

2.1.7. Selectin Crystal Structures

There are six crystal structures of selectins deposited on the protein data base:

- E-selectin lectin and EGF-like domain in absence of ligand (PDB code 1ESL)¹⁵.
- E-selectin lectin and EGF-like domain soaked with sLe^x (PDB code 1G1T)⁷⁹.
- E-selectin lectin and EGF-like domain soaked with quinic acid¹⁰⁰. The coordinates have not been deposited on the PDB. However the overall fold has been described to match previous E-selectin structures.
- P-selectin lectin and EGF-like domain in absence of ligand (PDB code 1G1R)⁷⁹.
- P-selectin lectin and EGF-like domain soaked with sLe^x (PDB code 1G1Q)⁷⁹.
- P-selectin lectin and EGF-like domain co-crystallized with SGP-3, an sLe^x bearing peptide fragment of PSGL-1 (PDB code 1G1S)⁷⁹.
- L-selectin lectin and EGF-like domain in absence of ligand (PDB code 3CFW).

The overall fold of all selectins in absence of ligand or soaked with sLe^x are similar and are in a bent conformation, with an extensive hydrogen network between the lectin and the EGF-like domain (Figure 2-7). However, the structure of P-selectin co-crystallized with SGP-3 (PDB code 1G1S)⁷⁹ shows a distinct overall conformation. SGP-3 is a peptide consisting of the first 19 N-terminal residues of PSGL-1, together with O-glycosylated Thr16 and three sulfated tyrosines (Tys6, Tys8, Tys11). The disruption of interdomain interactions in the P-selectin–SGP-3 structure leads to an extended conformation. This has also direct influence on the binding site and is thought to be a part of the allosteric mechanism involved in the catch-bond behavior of P-selectin. It is conceivable that this induced fit is not observed in soaked crystal structures, in which the protein is locked in preformed crystals.

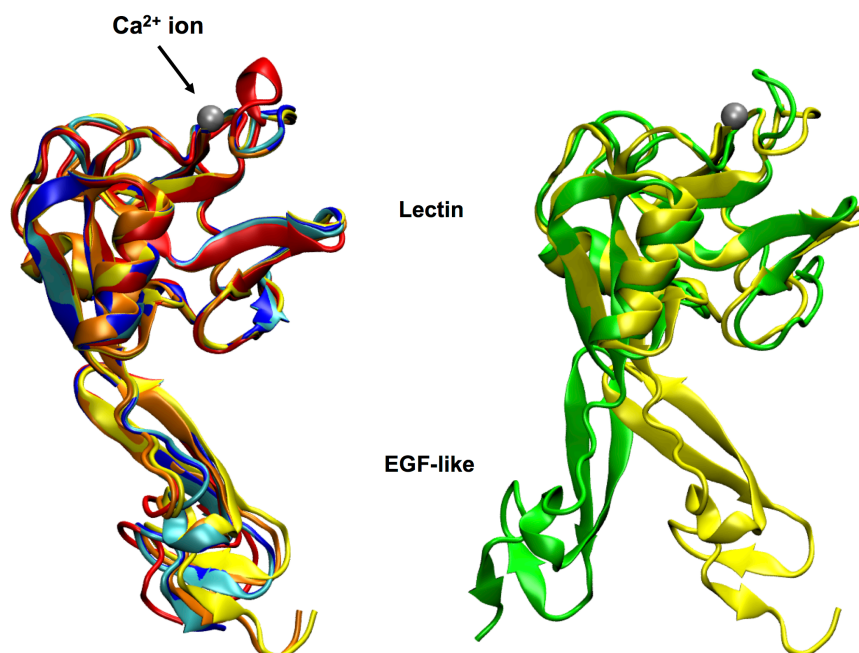


Figure 2-7. Superposition of available selectin structures. Left: The bent conformation as observed in apo- (blue, 1ESL)¹⁵ and sLe^x soaked (cyan, 1G1T)⁷⁹ E-selectin, apo- (yellow, 1G1Q)⁷⁹ and sLe^x soaked (orange, 1G1R)⁷⁹ P-selectin, and apo-L-selectin (red, 3CFW). Right: Comparison of apo-P-selectin (yellow) in the bent conformation with P-selectin co-crystallized with SGP-1, a PSGL-1 fragment (green, 1G1S). The Ca²⁺ ion is depicted as grey sphere. The ligands are not shown.

The following main observations between the bent and the extended conformation of P-selectin were described by Springer⁸⁰:

- Three loops in the lectin domain are altered, while a core part of the lectin domain remains unaltered.
- In the binding site, Asn83 and Asp89 exchange their position. Asp89 thereby replaces Asn83 as Ca²⁺ coordinating residue, and additionally makes a contact to the 2-hydroxyl position of sLe^x.
- Glu107 forms an additional interaction with the 2-hydroxyl group.
- The movement occurs around a pivot, which is formed by the Trp1 residue.
- The angle between the lectin domain and the EGF-like domain is opened by ~30°.

2.1.8. Selectin Antagonists

Two main strategies have been followed in the search for the treatment of selectin-related diseases. One approach is the application of biologicals, *i.e.* therapeutic antibodies or recombinant glycoproteins such as **GSP-6**¹⁰¹ (see Figure 2-8), which are not only specific for selectins, but also inhibit their function¹⁰²⁻¹⁰⁶. The major drawbacks of biologicals are their high production expenses and their requirement for intravenous application due to lack of oral bioavailability. Therefore, with a better understanding of the selectin–sLe^x interaction, efforts were subsequently focused on small-molecule selectin antagonists¹⁰⁷. For several drug development programs, sLe^x served as lead structure. The replacement of non-binding portions of sLe^x by hydrophobic parts lead to **CGP69669A**¹⁰⁸, which was further developed to the pan-selectin antagonist **GMI-1070**¹⁰⁹ which is currently in phase 2 clinical trials for vascular occlusion in sickle cell disease. Other developments lead to biphenyl substituted D-mannose, which yielded the pan-selectin antagonist **TBC1269**, a compound which is in phase 2 clinical trials for psoriasis and asthma¹¹⁰. Finally, non-carbohydrate ligands were developed such as the quinoline salicylic acid **PSI-697**¹¹¹, a P-selectin antagonists in phase I clinical trials.

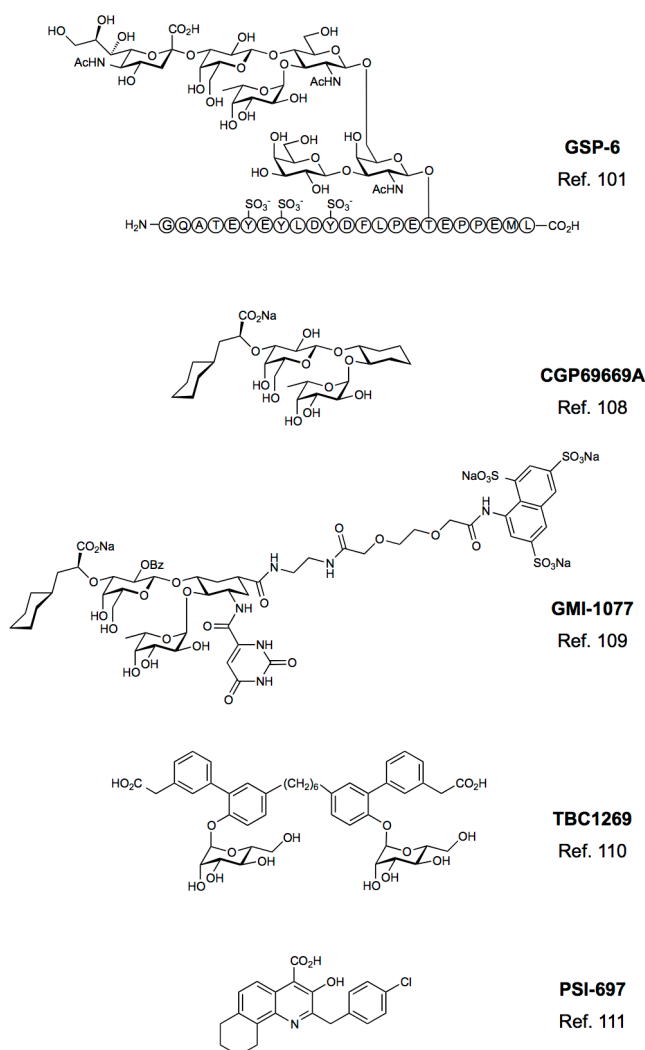


Figure 2-8. Representative list of selectin antagonists in advanced stages of drug discovery and development.

2.1.9. Recent Advances in Glycomimetic Antagonists for E-selectin

Although several selectin antagonists have entered clinical trials, no small-molecule selectin antagonist with high affinity has been identified to date. The IC_{50}/K_D values usually range between micromolar to millimolar affinity. The low affinity is mainly attributed to the shape of the binding site, as observed in the selectin–sLe^x crystal structures. It is shallow and solvent exposed and does therefore not allow for tight binding. Concerning sLe^x-based glycomimetic compounds for E-selectin, rational drug design has been mainly guided by NMR studies. Using transferred tr-NOE NMR, the bioactive conformation of sLe^x was identified and has been shown to be crucial for binding affinity of glycomimetics^{94,95}. In the example of **CGP69669A** with an IC_{50} value of $80\ \mu M$ ¹¹², this bioactive conformation of the relevant pharmacophores is not entirely fulfilled in solution. By systematic introduction of

substituents, the bioactive conformation was locked in **DS04-115**, which improved the affinity almost seven-fold compared to **CGP69669** (Figure 2-9). Another improvement in affinity was achieved in **GMI-1077**, which is one of the most potent E-selectin antagonist with a molecular weight below 1000 Da.

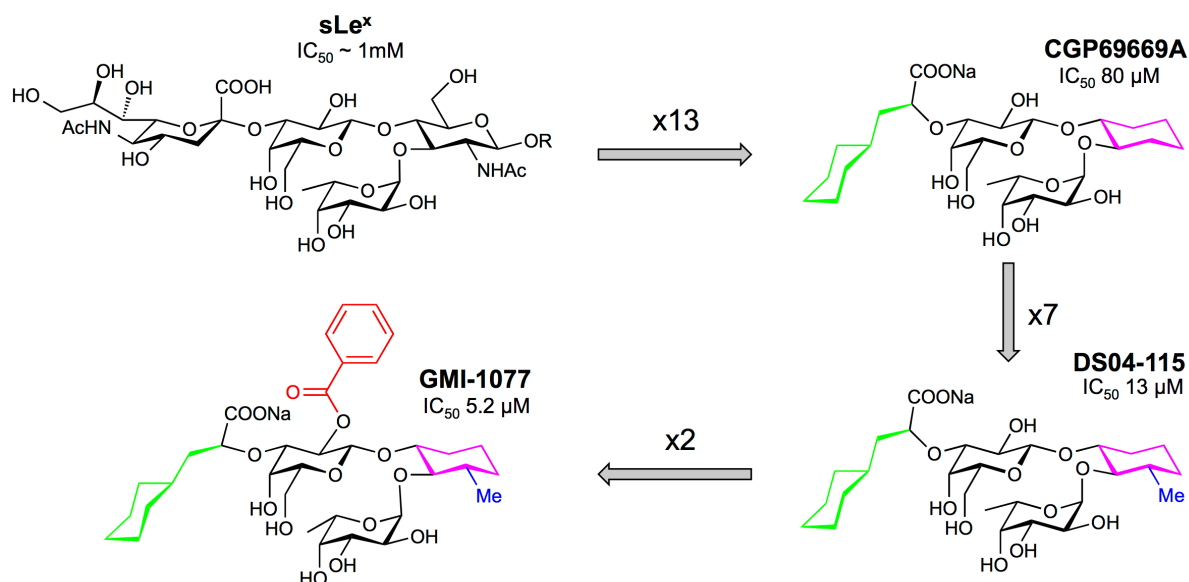


Figure 2-9. Recent advances in sLe^x glycomimetics as E-selectin antagonists. IC₅₀ values are taken from Schwizer *et al*¹¹².

2.1.10. References for Chapter 2.1.

1. Brodt, P., Fallavollita, L., Bresalier, R.S., Meterissian, S., Norton, C.R. & Wolitzky, B.A. Liver endothelial E-selectin mediates carcinoma cell adhesion and promotes liver metastasis. *Int. J. Cancer* **71**, 612-619 (1997).
2. Kansas, G.S., Saunders, K.B., Ley, K., Zakrzewicz, A., Gibson, R.M., Furie, B.C., Furie, B. & Tedder, T.F. A role for the epidermal growth factor-like domain of P-selectin in ligand recognition and cell adhesion. *J. Cell Biol.* **124**, 609-18 (1994).
3. Ley, K. & Kansas, G.S. Selectins in T-cell recruitment to non-lymphoid tissues and sites of inflammation. *Nat. Rev. Immunol.* **4**, 325-335 (2004).
4. Springer, T.A. Traffic Signals on Endothelium for Lymphocyte Recirculation and Leukocyte Emigration. *Annu. Rev. Physiol.* **57**, 827-872 (1995).
5. Lewinsohn, D.M., Bargatze, R.F. & Butcher, E.C. Leukocyte-Endothelial Cell Recognition - Evidence of a Common Molecular Mechanism Shared by Neutrophils, Lymphocytes, and Other Leukocytes. *J. Immunol.* **138**, 4313-4321 (1987).
6. Gallatin, W.M., Weissman, I.L. & Butcher, E.C. A cell-surface molecule involved in organ-specific homing of lymphocytes. *J. Immunol.* **177**, 5-9 (2006).
7. Siegelman, M.H., Vanderijn, M. & Weissman, I.L. Mouse Lymph-Node Homing Receptor Cdna Clone Encodes a Glycoprotein Revealing Tandem Interaction Domains. *Science* **243**, 1165-1172 (1989).
8. Gallatin, W.M., Weissman, I.L. & Butcher, E.C. A Cell-Surface Molecule Involved in Organ-Specific Homing of Lymphocytes. *Nature* **304**, 30-34 (1983).
9. Hsulín, S.C., Berman, C.L., Furie, B.C., August, D. & Furie, B. A Platelet Membrane-Protein Expressed during Platelet Activation and Secretion - Studies Using a Monoclonal-Antibody Specific for Thrombin-Activated Platelets. *J. Biol. Chem.* **259**, 9121-9126 (1984).
10. Mcever, R.P., Beckstead, J.H., Moore, K.L., Marshallcarlson, L. & Bainton, D.F. Gmp-140, a Platelet Alpha-Granule Membrane-Protein, Is Also Synthesized by Vascular Endothelial-Cells and Is Localized in Weibel-Palade Bodies. *J. Clin. Invest.* **84**, 92-99 (1989).
11. Palabrica, T., Lobb, R., Furie, B.C., Aronovitz, M., Benjamin, C., Hsu, Y.M., Sajer, S.A. & Furie, B. Leukocyte Accumulation Promoting Fibrin Deposition Is Mediated In vivo by P-Selectin on Adherent Platelets. *Nature* **359**, 848-851 (1992).
12. Erbe, D.V., Wolitzky, B.A., Presta, L.G., Norton, C.R., Ramos, R.J., Burns, D.K., Rumberger, J.M., Rao, B.N.N., Foxall, C., Brandley, B.K. et al. Identification of an E-selectin region critical for carbohydrate recognition and cell-adhesion. *J. Cell Biol.* **119**, 215-227 (1992).
13. Pigott, R., Needham, L.A., Edwards, R.M., Walker, C. & Power, C. Structural and Functional-Studies of the Endothelial Activation Antigen Endothelial Leukocyte Adhesion Molecule-1 Using a Panel of Monoclonal-Antibodies. *J. Immunol.* **147**, 130-135 (1991).
14. Bowen, B.R., Fennie, C. & Lasky, L.A. The Mel-14 Antibody Binds to the Lectin Domain of the Murine Peripheral Lymph-Node Homing Receptor. *J. Cell Biol.* **110**, 147-153 (1990).
15. Graves, B.J., Crowther, R.L., Chandran, C., Rumberger, J.M., Li, S., Huang, K.S., Presky, D.H., Familletti, P.C., Wolitzky, B.A. & Burns, D.K. Insight into E-selectin ligand interaction from the crystal-structure and mutagenesis of the LEC EGF domains. *Nature* **367**, 532-538 (1994).
16. Vestweber, D. & Blanks, J.E. Mechanisms that regulate the function of the selectins and their ligands. *Physiol. Rev.* **79**, 181-213 (1999).

17. Li, S.H., Burns, D.K., Rumberger, J.M., Presky, D.H., Wilkinson, V.L., Anostario, M., Jr., Wolitzky, B.A., Norton, C.R., Familletti, P.C., Kim, K.J. et al. Consensus repeat domains of E-selectin enhance ligand binding. *J. Biol. Chem.* **269**, 4431-4437 (1994).
18. Tu, L., Chen, A., Delahunty, M.D., Moore, K.L., Watson, S.R., McEver, R.P. & Tedder, T.F. L-selectin binds to P-selectin glycoprotein ligand-1 on leukocytes: interactions between the lectin, epidermal growth factor, and consensus repeat domains of the selectins determine ligand binding specificity. *J. Immunol.* **157**, 3995-4004 (1996).
19. Watson, S.R., Imai, Y., Fennie, C., Geoffrey, J., Singer, M., Rosen, S.D. & Lasky, L.A. The Complement Binding-Like Domains of the Murine Homing Receptor Facilitate Lectin Activity. *J. Cell Biol.* **115**, 235-243 (1991).
20. Kolbinger, F., Patton, J.T., Geisenhoff, G., Aenis, A., Li, X.H. & Katopodis, A.G. The carbohydrate-recognition domain of E-selectin is sufficient for ligand binding under both static and flow conditions. *Biochemistry (Mosc.)* **35**, 6385-6392 (1996).
21. Sarangapani, K.K., Marshall, B.T., McEver, R.P. & Zhu, C. Molecular stiffness of selectins. *J. Biol. Chem.* **286**, 9567-76 (2011).
22. Patel, K.D., Nollert, M.U. & McEver, R.P. P-selectin must extend a sufficient length from the plasma membrane to mediate rolling of neutrophils. *J. Cell Biol.* **131**, 1893-1902 (1995).
23. Lorenzon, P., Vecile, E., Nardon, E., Ferrero, E., Harlan, J.M., Tedesco, F. & Dobrina, A. Endothelial cell E- and P-selectin and vascular cell adhesion molecule-1 function as signaling receptors. *J. Cell Biol.* **142**, 1381-1391 (1998).
24. Hu, Y.Y., Szente, B., Kiely, J.M. & Gimbrone, M.A. Molecular events in transmembrane signaling via E-selectin - SHP2 association, adaptor protein complex formation and ERK1/2 activation. *J. Biol. Chem.* **276**, 48549-48553 (2001).
25. Yoshida, M., Westlin, W.F., Wang, N., Ingber, D.E., Rosenzweig, A., Resnick, N. & Gimbrone, M.A. Leukocyte adhesion to vascular endothelium induces E-selectin linkage to the actin cytoskeleton. *J. Cell Biol.* **133**, 445-455 (1996).
26. Phillips, M.L., Nudelman, E., Gaeta, F.C.A., Perez, M., Singhal, A.K., Hakomori, S.I. & Paulson, J.C. Elam-1 Mediates Cell-Adhesion by Recognition of a Carbohydrate Ligand, Sialyl-Lex. *Science* **250**, 1130-1132 (1990).
27. Ohmori, K., Kanda, K., Mitsuoka, C., Kanamori, A., Kurata-Miura, K., Sasaki, K., Nishi, T., Tamatani, T. & Kannagi, R. P- and E-selectins recognize sialyl 6-sulfo Lewis X, the recently identified L-selectin ligand. *Biochem. Biophys. Res. Commun.* **278**, 90-96 (2000).
28. Poppe, L., Brown, G.S., Philo, J.S., Nikrad, P.V. & Shah, B.H. Conformation of sLe(x) tetrasaccharide, free in solution and bound to E-, P-, and L-selectin. *J. Am. Chem. Soc.* **119**, 1727-1736 (1997).
29. McEver, R.P. P-selectin and PSGL-1: Exploiting connections between inflammation and venous thrombosis. *Thromb. Haemost.* **87**, 364-365 (2002).
30. Mehta, P., Cummings, R.D. & McEver, R.P. Affinity and kinetic analysis of P-selectin binding to P-selectin glycoprotein ligand-1. *J. Biol. Chem.* **273**, 32506-32513 (1998).
31. Li, F.G., Wilkins, P.P., Crawley, S., Weinstein, J., Cummings, R.D. & McEver, R.P. Post-translational modifications of recombinant P-selectin glycoprotein ligand-1 required for binding to P- and E-selectin. *J. Biol. Chem.* **271**, 3255-3264 (1996).
32. Lenter, M., Levinovitz, A., Isenmann, S. & Vestweber, D. Monospecific and Common Glycoprotein Ligands for E-Selectin and P-Selectin on Myeloid Cells. *J. Cell Biol.* **125**, 471-481 (1994).

33. Levinovitz, A., Muhlhoff, J., Isenmann, S. & Vestweber, D. Identification of a glycoprotein ligand for E-selectin on mouse myeloid cells. *J. Cell Biol.* **121**, 449-459 (1993).
34. Wild, M.K., Huang, M.C., Schulze-Horsel, U., van der Merwe, P.A. & Vestweber, D. Affinity, kinetics, and thermodynamics of E-selectin binding to E-selectin ligand-1. *J. Biol. Chem.* **276**, 31602-31612 (2001).
35. Streeter, P.R., Berg, E.L., Rouse, B.T.N., Bargatze, R.F. & Butcher, E.C. A Tissue-Specific Endothelial-Cell Molecule Involved in Lymphocyte Homing. *Nature* **331**, 41-46 (1988).
36. Nakache, M., Berg, E.L., Streeter, P.R. & Butcher, E.C. The Mucosal Vascular Addressin Is a Tissue-Specific Endothelial-Cell Adhesion Molecule for Circulating Lymphocytes. *Nature* **337**, 179-181 (1989).
37. Baumhueter, S., Singer, M.S., Henzel, W., Hemmerich, S., Renz, M., Rosen, S.D. & Lasky, L.A. Binding of L-Selectin to the Vascular Sialomucin Cd34. *Science* **262**, 436-438 (1993).
38. Kanda, H., Tanaka, T., Matsumoto, M., Umemoto, E., Ebisuno, Y., Kinoshita, M., Noda, M., Kannagi, R., Hirata, T., Murai, T. et al. Endomucin, a sialomucin expressed in high endothelial venules, supports L-selectin-mediated rolling. *Int. Immunol.* **16**, 1265-1274 (2004).
39. Fieger, C.B., Sassetti, C.M. & Rosen, S.D. Endoglycan, a member of the CD34 family, functions as an L-selectin ligand through modification with tyrosine sulfation and sialyl Lewis x. *J. Biol. Chem.* **278**, 27390-27398 (2003).
40. Imai, Y., Singer, M.S., Fennie, C., Lasky, L.A. & Rosen, S.D. Identification of a Carbohydrate-Based Endothelial Ligand for a Lymphocyte Homing Receptor. *J. Cell Biol.* **113**, 1213-1221 (1991).
41. Geng, J.G., Bevilacqua, M.P., Moore, K.L., McIntyre, T.M., Prescott, S.M., Kim, J.M., Bliss, G.A., Zimmerman, G.A. & McEver, R.P. Rapid Neutrophil Adhesion to Activated Endothelium Mediated by Gmp-140. *Nature* **343**, 757-760 (1990).
42. Hattori, R., Hamilton, K.K., Fugate, R.D., McEver, R.P. & Sims, P.J. Stimulated Secretion of Endothelial VonWillebrand-Factor Is Accompanied by Rapid Redistribution to the Cell-Surface of the Intracellular Granule Membrane-Protein Gmp-140. *J. Biol. Chem.* **264**, 7768-7771 (1989).
43. Modur, V., Feldhaus, M.J., Weyrich, A.S., Jicha, D.L., Prescott, S.M., Zimmerman, G.A. & McIntyre, T.M. Oncostatin M is a proinflammatory mediator - In vivo effects correlate with endothelial cell expression of inflammatory cytokines and adhesion molecules. *J. Clin. Invest.* **100**, 158-168 (1997).
44. Yao, L.B., Pan, J.L., Setiadi, H., Patel, K.D. & McEver, R.P. Interleukin 4 or oncostatin M induces a prolonged increase in P-selectin mRNA and protein in human endothelial cells. *J. Exp. Med.* **184**, 81-92 (1996).
45. Woltmann, G., McNulty, C.A., Dewson, G., Symon, F.A. & Wardlaw, A.J. Interleukin-13 induces PSGL-1/P-selectin-dependent adhesion of eosinophils, but not neutrophils, to human umbilical vein endothelial cells under flow. *Blood* **95**, 3146-3152 (2000).
46. Bevilacqua, M.P., Pober, J.S., Mendrick, D.L., Cotran, R.S. & Gimbrone, M.A., Jr. Identification of an inducible endothelial-leukocyte adhesion molecule. *Proc. Natl. Acad. Sci. U. S. A.* **84**, 9238-42 (1987).
47. Wyble, C.W., Hynes, K.L., Kuchibhotla, J., Marcus, B.C., Hallahan, D. & Gewertz, B.L. TNF-alpha and IL-1 upregulate membrane-bound and soluble E-selectin through a common pathway. *J. Surg. Res.* **73**, 107-112 (1997).

48. Bullard, D.C., Kunkel, E.J., Kubo, H., Hicks, M.J., Lorenzo, I., Doyle, N.A., Doerschuk, C.M., Ley, K. & Beaudet, A.L. Infectious susceptibility and severe deficiency of leukocyte rolling and recruitment in E-selectin and P-selectin double mutant mice. *J. Exp. Med.* **183**, 2329-2336 (1996).
49. Mitchell, D.J., Li, P., Reinhardt, P.H. & Kubes, P. Importance of L-selectin-dependent leukocyte-leukocyte interactions in human whole blood. *Blood* **95**, 2954-2959 (2000).
50. Briaud, S.A., Ding, Z.M., Michael, L.H., Entman, M.L., Daniel, S. & Ballantyne, C.M. Leukocyte trafficking and myocardial reperfusion injury in ICAM-1/P-selectin-knockout mice. *Am J Physiol-Heart C* **280**, H60-H67 (2001).
51. Dejana, E. The transcellular railway: insights into leukocyte diapedesis. *Nat. Cell Biol.* **8**, 105-107 (2006).
52. Ulbrich, H., Eriksson, E.E. & Lindbom, L. Leukocyte and endothelial cell adhesion molecules as targets for therapeutic interventions in inflammatory disease. *Trends Pharmacol. Sci.* **24**, 640-647 (2003).
53. Ross, R. The Pathogenesis of Atherosclerosis - a Perspective for the 1990s. *Nature* **362**, 801-809 (1993).
54. Pellegatta, F., Pizzetti, G., Lu, Y., Radaelli, A., Pomes, D., Carlino, M., Meloni, C., Belotti, G., Galli, L., Vidal, M.J. et al. Soluble E-selectin and intercellular adhesion molecule-1 plasma levels increase during acute myocardial infarction. *J. Cardiovasc. Pharmacol.* **30**, 455-460 (1997).
55. Shreeniwas, R., Koga, S., Karakurum, M., Pinsky, D., Kaiser, E., Brett, J., Wolitzky, B.A., Norton, C., Plocinski, J., Benjamin, W. et al. Hypoxia-Mediated Induction of Endothelial-Cell Interleukin-1-Alpha - an Autocrine Mechanism Promoting Expression of Leukocyte Adhesion Molecules on the Vessel Surface. *J. Clin. Invest.* **90**, 2333-2339 (1992).
56. Billups, K.L., Palladino, M.A., Hinton, B.T. & Sherley, J.L. Expression of E-Selectin Messenger-Rna during Ischemia-Reperfusion Injury. *J. Lab. Clin. Med.* **125**, 626-633 (1995).
57. Palluy, O., Morliere, L., Gris, J.C., Bonne, C. & Modat, G. Hypoxia Reoxygenation Stimulates Endothelium to Promote Neutrophil Adhesion. *Free Radic. Biol. Med.* **13**, 21-30 (1992).
58. De Caterina, R., Ghiadoni, L., Taddei, S., Virdis, A., Almerigogna, F., Basta, G., Lazzarini, G., Bernini, W. & Salvetti, A. Soluble E-selectin in essential hypertension: A correlate of vascular structural changes. *Am. J. Hypertens.* **14**, 259-266 (2001).
59. Albertini, J.P., Valensi, P., Lormeau, B., Aourousseau, M.H., Ferriere, F., Attali, J.R. & Gattegno, L. Elevated concentrations of soluble E-selectin and vascular cell adhesion molecule-1 in NIDDM - Effect of intensive insulin treatment. *Diabetes Care* **21**, 1008-1013 (1998).
60. Cronstein, B.N. & Weissmann, G. The Adhesion Molecules of Inflammation. *Arthritis Rheum.* **36**, 147-157 (1993).
61. Spronk, P.E., Bootsma, H., Huitema, M.G., Limburg, P.C. & Kallenberg, C.G.M. Levels of Soluble Vcam-1, Soluble Icam-1, and Soluble E-Selectin during Disease Exacerbations in Patients with Systemic Lupus-Erythematosus (Sle) - a Long-Term Prospective-Study. *Clin. Exp. Immunol.* **97**, 439-444 (1994).
62. Gundel, R.H., Wegner, C.D., Torcellini, C.A., Clarke, C.C., Haynes, N., Rothlein, R., Smith, C.W. & Letts, L.G. Endothelial Leukocyte Adhesion Molecule-1 Mediates Antigen-Induced Acute Airway Inflammation and Late-Phase Airway-Obstruction in Monkeys. *J. Clin. Invest.* **88**, 1407-1411 (1991).

63. Matsui, N.M., Borsig, L., Rosen, S.D., Yaghmai, M., Varki, A. & Embury, S.H. P-selectin mediates the adhesion of sickle erythrocytes to the endothelium. *Blood* **98**, 1955-1962 (2001).
64. Yoshida, M., Takano, Y., Sasaoka, T., Izumi, T. & Kimura, A. E-selectin polymorphism associated with myocardial infarction causes enhanced leukocyte-endothelial interactions under flow conditions. *Arterioscler. Thromb. Vasc. Biol.* **23**, 783-788 (2003).
65. Nadi, E., Hajilooii, M., Zeraati, F., Ansari, M., Tavana, S., Hashemi, S.H. & Rafiei, A. E-selectin S128R polymorphism leads to severe asthma. *Iran J Allergy Asthm* **6**, 49-57 (2007).
66. Alessandro, R., Seidita, G., Flugy, A.M., Damiani, F., Russo, A., Corrado, C., Colomba, P., Gullotti, L., Buettner, R., Bruno, L. et al. Role of S128R polymorphism of E-selectin in colon metastasis formation. *Int. J. Cancer* **121**, 528-535 (2007).
67. Wenzel, K., Ernst, M., Rohde, K., Baumann, G. & Speer, A. DNA polymorphisms in adhesion molecule genes - A new risk factor for early atherosclerosis. *Hum. Genet.* **97**, 15-20 (1996).
68. Marteau, J.B., Sass, C., Pfister, M., Lambert, D., Noyer-Weidner, M. & Visvikis, S. The Leu554Phe polymorphism in the E-selectin gene is associated with blood pressure in overweight people. *J. Hypertens.* **22**, 305-311 (2004).
69. Bhatia, S.K., King, M.R. & Hammer, D.A. The state diagram for cell adhesion mediated by two receptors. *Biophys. J.* **84**, 2671-2690 (2003).
70. Lawrence, M.B. & Springer, T.A. Leukocytes Roll on a Selectin at Physiological Flow-Rates - Distinction from and Prerequisite for Adhesion through Integrins. *Cell* **65**, 859-873 (1991).
71. Brunk, D.K. & Hammer, D.A. Quantifying rolling adhesion with a cell-free assay: E-selectin and its carbohydrate ligands. *Biophys. J.* **72**, 2820-2833 (1997).
72. Hammer, D.A. & Apte, S.M. Simulation of Cell Rolling and Adhesion on Surfaces in Shear-Flow - General Results and Analysis of Selectin-Mediated Neutrophil Adhesion. *Biophys. J.* **63**, 35-57 (1992).
73. Hammer, D.A., Tempelman, L.A. & Apte, S.M. Statistics of Cell-Adhesion under Hydrodynamic Flow - Simulation and Experiment. *Blood Cells* **19**, 261-277 (1993).
74. Kansas, G.S. Selectins and their ligands: Current concepts and controversies. *Blood* **88**, 3259-3287 (1996).
75. Lawrence, M.B., Kansas, G.S., Kunkel, E.J. & Ley, K. Threshold levels of fluid shear promote leukocyte adhesion through selectins (CD62L,P,E). *FASEB J.* **11**, 1755-1755 (1997).
76. Marshall, B.T., Long, M., Piper, J.W., Yago, T., McEver, R.P. & Zhu, C. Direct observation of catch bonds involving cell-adhesion molecules. *Nature* **423**, 190-193 (2003).
77. Wayman, A.M., Chen, W., McEver, R.P. & Zhu, C. Triphasic force dependence of E-selectin/ligand dissociation governs cell rolling under flow. *Biophys. J.* **99**, 1166-1174 (2010).
78. Zhu, C., Yago, T., Lou, J., Zarnitsyna, V.I. & McEver, R.P. Mechanisms for flow-enhanced cell adhesion. *Ann. Biomed. Eng.* **36**, 604-21 (2008).
79. Somers, W.S., Tang, J., Shaw, G.D. & Camphausen, R.T. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe(X) and PSGL-1. *Cell* **103**, 467-479 (2000).
80. Springer, T.A. Structural basis for selectin mechanochemistry. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 91-96 (2009).

81. Phan, U.T., Waldron, T.T. & Springer, T.A. Remodeling of the lectin-EGF-like domain interface in P- and L-selectin increases adhesiveness and shear resistance under hydrodynamic force. *Nat. Immunol.* **7**, 883-889 (2006).
82. Waldron, T.T. & Springer, T.A. Transmission of allostery through the lectin domain in selectin-mediated cell adhesion. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 85-90 (2009).
83. Lou, J.Z., Yago, T., Klopocki, A.G., Mehta, P., Chen, W., Zarnitsyna, V.I., Bovin, N.V., Zhu, C. & McEver, R.P. Flow-enhanced adhesion regulated by a selectin interdomain hinge. *J. Cell Biol.* **174**, 1107-1117 (2006).
84. Ramphal, J.Y., Zheng, Z.L., Perez, C., Walker, L.E., Defrees, S.A. & Gaeta, F.C.A. Structure-Activity-Relationships of Sialyl-Lewis X-Containing Oligosaccharides .1. Effect of Modifications of the Fucose Moiety. *J. Med. Chem.* **37**, 3459-3463 (1994).
85. Brandley, B.K., Kiso, M., Abbas, S., Nikrad, P., Srivasatava, O., Foxall, C., Oda, Y. & Hasegawa, A. Structure-Function Studies on Selectin Carbohydrate Ligands - Modifications to Fucose, Sialic-Acid and Sulfate as a Sialic-Acid Replacement. *Glycobiology* **3**, 633-641 (1993).
86. Stahl, W., Sprengard, U., Kretzschmar, G. & Kunz, H. Synthesis of Deoxy Sialyl Lewis(X) Analogs, Potential Selectin Antagonists. *Angew. Chem. Int. Ed. Engl.* **33**, 2096-2098 (1994).
87. Banteli, R. & Ernst, B. Synthesis of sialyl Lewis(X) mimics, modifications of the 6-position of galactose. *Bioorg. Med. Chem. Lett.* **11**, 459-462 (2001).
88. Komba, S., Ishida, H., Kiso, M. & Hasegawa, A. Synthesis of deoxygalactose-containing sialyl Le(X) ganglioside analogues to elucidate the structure necessary for selectin recognition. *Glycoconj. J.* **13**, 241-254 (1996).
89. Tyrrell, D., James, P., Rao, N., Foxall, C., Abbas, S., Dasgupta, F., Nashed, M., Hasegawa, A., Kiso, M., Asa, D. et al. Structural Requirements for the Carbohydrate Ligand of E-Selectin. *Proc. Natl. Acad. Sci. U. S. A.* **88**, 10372-10376 (1991).
90. Defrees, S.A., Gaeta, F.C.A., Lin, Y.C., Ichikawa, Y. & Wong, C.H. Ligand Recognition by E-Selectin - Analysis of Conformation and Activity of Synthetic Monomeric and Bivalent Sialyl Lewis-X Analogs. *J. Am. Chem. Soc.* **115**, 7549-7550 (1993).
91. Wada, Y., Saito, T., Matsuda, N., Ohmoto, H., Yoshino, K., Ohashi, M., Kondo, H., Ishida, H., Kiso, M. & Hasegawa, A. Studies on selectin blockers .2. Novel selectin blocker as potential therapeutics for inflammatory disorders. *J. Med. Chem.* **39**, 2055-2059 (1996).
92. Thoma, G., Magnani, J.L., Patton, J.T., Ernst, B. & Jahnke, W. Preorganization of the bioactive conformation of sialyl Lewis(X) analogues correlates with their affinity to E-selectin. *Angew. Chem. Int. Ed.* **40**, 1941-1945 (2001).
93. Scheffler, K., Ernst, B., Katopodis, A., Magnani, J.L., Wang, W.T., Weisemann, R. & Peters, T. Determination of the bioactive conformation of the carbohydrate ligand in the E-selectin sialyl Lewis(x) complex. *Angew. Chem. Int. Ed. Engl.* **34**, 1841-1844 (1995).
94. Harris, R., Kiddle, G.R., Field, R.A., Milton, M.J., Ernst, B., Magnani, J.L. & Homans, S.W. Stable-isotope-assisted NMR studies on C-13-enriched sialyl Lewis(x) in solution and bound to E-selectin. *J. Am. Chem. Soc.* **121**, 2546-2551 (1999).
95. Scheffler, K., Brisson, J.R., Weisemann, R., Magnani, J.L., Wong, W.T., Ernst, B. & Peters, T. Application of homonuclear 3D NMR experiments and 1D analogs to study the conformation of sialyl Lewis(x) bound to E-selectin. *J. Biomol. NMR* **9**, 423-436 (1997).

96. Rutherford, T.J., Spackman, D.G., Simpson, P.J. & Homans, S.W. 5 Nanosecond Molecular-Dynamics and Nmr-Study of Conformational Transitions in the Sialyl-Lewis-X Antigen. *Glycobiology* **4**, 59-68 (1994).
97. Ichikawa, Y., Lin, Y.C., Dumas, D.P., Shen, G.J., Garciajunceda, E., Williams, M.A., Bayer, R., Ketcham, C., Walker, L.E., Paulson, J.C. et al. Chemical-Enzymatic Synthesis and Conformational-Analysis of Sialyl Lewis-X and Derivatives. *J. Am. Chem. Soc.* **114**, 9283-9298 (1992).
98. Lin, Y.C., Hummel, C.W., Huang, D.H., Ichikawa, Y., Nicolaou, K.C. & Wong, C.H. Conformational Studies of Sialyl Lewis-X in Aqueous-Solution. *J. Am. Chem. Soc.* **114**, 5452-5454 (1992).
99. Chervenak, M.C. & Toone, E.J. Calorimetric Analysis of the Binding of Lectins with Overlapping Carbohydrate-Binding Ligand Specificities. *Biochemistry (Mosc.)* **34**, 5685-5695 (1995).
100. Kaila, N., Somers, W.S., Thomas, B.E., Thakker, P., Janz, K., DeBernardo, S., Tam, S., Moore, W.J., Yang, R.Y., Wrona, W. et al. Quinic acid derivatives as sialyl Lewis(x)-mimicking selectin inhibitors: Design, synthesis, and crystal structure in complex with E-selectin. *J. Med. Chem.* **48**, 4346-4357 (2005).
101. Leppanen, A., Mehta, P., Ouyang, Y.B., Ju, T.Z., Helin, J., Moore, K.L., van Die, I., Canfield, W.M., McEver, R.P. & Cummings, R.D. A novel glycosulfopeptide binds to P-selectin and inhibits leukocyte adhesion to P-selectin. *J. Biol. Chem.* **274**, 24838-24848 (1999).
102. Friedman, G., Jankowski, S., Shahla, M., Goldman, M., Rose, R.M., Kahn, R.J. & Vincent, J.L. Administration of an antibody to E-selectin in patients with septic shock. *Crit. Care Med.* **24**, 229-233 (1996).
103. Seekamp, A., van Griensven, M., Dhont, E., Diefenbeck, M., Demeyer, I., Vundelinckx, G., Haas, N., Schaechinger, U., Wolowickwa, L., Rammelt, S. et al. The effect of anti-L-selectin (aselizumab) in multiple traumatized patients - Results of a phase II clinical trial. *Crit. Care Med.* **32**, 2021-2028 (2004).
104. Bednar, M.M., Gross, C.E., Russell, S.R., Fuller, S.P., Ellenberger, C.L., Schindler, E., Klingbeil, C. & Vexler, V. Humanized anti-L-selectin monoclonal antibody DREG200 therapy in acute thromboembolic stroke. *Neurol. Res.* **20**, 403-408 (1998).
105. Suzuki, K., Fukushima, S., Coppen, S.R., Yamahara, K., Varela-Carver, A., Ermakov, A. & Yacoub, M.H. A novel strategy for myocardial protection by combined antibody therapy inhibiting both P-selectin and intercellular adhesion molecule-1 via retrograde intracoronary route. *J. Mol. Cell. Cardiol.* **40**, 975-975 (2006).
106. Ni, X., Gu, Y., Legos, J., Eppihimer, M., Schaub, R. & Tuma, R. Modulation of leukocyte-endothelial interactions following CNS ischemia. *7th World Congress for Microcirculation*, 53-58 (2001).
107. Bedard, P.W. & Kaila, N. Selectin inhibitors: a patent review. *Expert Opin Ther Pat* **20**, 781-793 (2010).
108. Norman, K.E., Anderson, G.P., Kolb, H.C., Ley, K. & Ernst, B. Sialyl Lewis(x) (sLe(x)) and an sLe(x) mimetic, CGP69669A, disrupt E-selectin-dependent leukocyte rolling in vivo. *Blood* **91**, 475-483 (1998).
109. Chang, J.S., Patton, J.T., Sarkar, A., Ernst, B., Magnani, J.L. & Frenette, P.S. GMI-1070, a novel pan-selectin antagonist, reverses acute vascular occlusions in sickle cell mice. *Blood* **116**, 1779-1786 (2010).
110. Avila, P.C., Boushey, H.A., Wong, H., Grundland, H., Liu, J. & Fahy, J.V. Effect of a single dose of the selectin inhibitor TBC1269 on early and late asthmatic responses. *Clin. Exp. Allergy* **34**, 77-84 (2004).

111. Kaila, N., Janz, K., Huang, A., Moretto, A., DeBernardo, S., Bedard, P.W., Tam, S., Clerin, V., Keith, J.C., Tsao, D.H.H. et al. 2-(4-chlorobenzyl)-3-hydroxy-7,8,9,10-tetrahydrobenzo[H]quinoline-4-carboxylic acid (PSI-697): Identification of a clinical candidate from the quinoline salicylic acid series of P-selectin antagonists. *J. Med. Chem.* **50**, 40-64 (2007).
112. Schwizer, D., Patton, J.T., Cutting, B., Smiesko, M., Wagner, B., Kato, A., Weckerle, C., Binder, F.P., Rabbani, S., Schwardt, O. et al. Pre-organization of the core structure of E-selectin antagonists. *Chemistry* **18**, 1342-1351 (2012).

Chapter 2.2. – Publication 1

Sialyl Lewis^x: A "Pre-Organized Water Oligomer"?

Published in *Angewandte Chemie Int. Ed.* **51**(29): 7327-7331.

Contributions

- | | |
|---------------|--|
| F.P.C. Binder | <ul style="list-style-type: none">• Compound synthesis.• Manuscript preparation. |
| K. Lemme | <ul style="list-style-type: none">• Protein purification.• Isothermal Titration Calorimetry.• Manuscript preparation. |
| R. C. Preston | <ul style="list-style-type: none">• Generation of functional antibody column for purification.• Aiding in manuscript preparation. |

Sialyl Lewis^x: A “Pre-Organized Water Oligomer”?**

Florian P. C. Binder, Katrin Lemme, Roland C. Preston, and Beat Ernst*

In memory of Daniel Bellùs

In recent years, lectins, such as selectins,^[1] galectins,^[2] or siglecs^[3] have received increasing attention as drug targets. Among them, selectins are the most extensively studied, since they are key players in the early stages of inflammation and therefore promising targets for the treatment of diseases with an inflammatory component, such as stroke, asthma, psoriasis, or rheumatoid arthritis.^[4] The key role of selectins is to promote the initial step of the inflammatory cascade, by allowing leukocytes to roll along the vascular endothelial surface. This step is followed by the integrin-mediated firm adhesion and the final extravasation to the site of the inflammatory stimulus.^[5]

The specific interaction between E-selectin and its physiological ligand ESL-1 (E-selectin ligand-1) is mediated by the tetrasaccharide sialyl Lewis^x (sLe^x, **1**).^[6] Consequently, sLe^x (**1**) became the lead structure for the search of drug-like, high-affinity selectin antagonists.^[1,7] Elucidation of the structure activity relationship (SAR),^[8] mutation studies,^[9] transferred nuclear overhauser enhancement NMR spectroscopy (trNOE-NMR),^[10] saturation transfer difference NMR spectroscopy (STD-NMR),^[11] molecular modeling,^[12] and finally X-ray crystallography^[13] yielded a precise picture of the interactions of sLe^x and E-selectin on an atomic level (Figure 1). Because docking studies^[7] and STD-NMR experiments^[11] revealed that the *N*-acetyl-D-glucosamine (D-GlcNAc) and *N*-acetyl-D-neuraminic acid (D-Neu5Ac) moieties have only weak interactions with the protein,^[12] they were replaced with structurally simplified mimics, resulting in E-selectin antagonists with significantly improved binding affinities. However, the affinity of these antagonists,

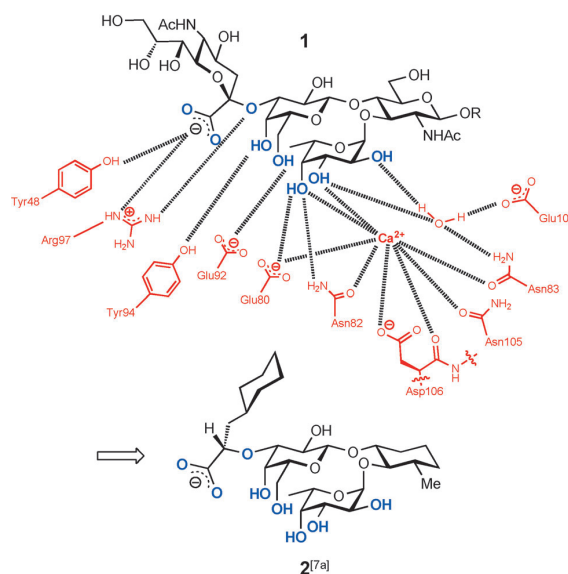


Figure 1. Top: Detailed representation of the interactions between sLe^x (**1**) and E-selectin as observed in the crystal structure;^[13] the pharmacophores of **1** are highlighted in blue. Bottom: The glycomimetic **2** exhibits a 13 μM affinity in a cell-free ligand-based competitive binding assay.^[7a]

for example, **2** in Figure 1,^[7a] is still only in the low micromolar range.

Despite the progress made, the driving force of the interaction of E-selectin with its ligands has not been fully characterized to date, neither for sLe^x (**1**), nor for any low molecular weight selectin antagonist. However, Wild et al. estimated the enthalpic contribution of the E-selectin/ESL-1 interaction by van't Hoff analysis, which involves the correlation of the binding affinity measured at different temperatures. The results indicated that enthalpic changes contribute only 10 to 25 % of the binding free energy ΔG and that the interaction is primarily driven by favorable entropy changes.^[14]

Recently, the thermodynamic aspects of protein–ligand interactions have gained increasing interest in drug discovery.^[15] Particularly enthalpy and entropy changes provide valuable information for lead optimization. Having access to these individual components of binding affinity rather than the overall value facilitates the successful design of high-affinity ligands. Herein, we report a comprehensive study on the thermodynamic fingerprint of a series of E-selectin antagonists.

[*] Dr. F. P. C. Binder,^[1] Dr. K. Lemme,^[1] R. C. Preston, Prof. Dr. B. Ernst
Institute of Molecular Pharmacy, University of Basel
Klingelbergstrasse 50, 4056 Basel (Switzerland)
E-mail: beat.ernst@unibas.ch

[†] These authors contributed equally to this work.

[**] We gratefully acknowledge the financial support by the Swiss National Science Foundation (grant no. 200020-103875/1) and by GlycoMimetics Inc., Gaithersburg, MD (USA). We are greatly indebted to Dr. Francis Bitsch and Peggy Brunet-LeFeuvre (Novartis, Basel, Switzerland) for giving us access to their VP-ITC as well as for their advice regarding experimental setup and data analysis. Finally, we are thankful to Dr. Martin Smiesko (Institute of Molecular Pharmacy, University of Basel) for preparing the sLe^x/E-selectin illustrations.

Supporting information for this article (details of the synthesis of antagonist **4**, the expression and purification of E-selectin/IgG, the competitive binding assay, and the isothermal calorimetry experiments) is available on the WWW under <http://dx.doi.org/10.1002/anie.201202555>.

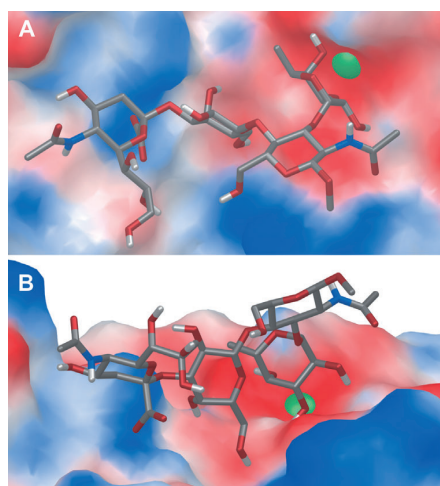


Figure 2. sLe^x (**1**) bound to E-selectin as observed in the crystal structure (protein data bank (PDB) code: 1G1T).^[13a] A) top view: the binding epitope on E-selectin is dominated by polar residues [polar residues in red (positively charged) and blue (negatively charged), nonpolar in white]. The contact area is 275 Å². B) side view: Only a small part of sLe^x directly contributes to binding (graphics generated by Maestro^[32]).

a scaffold to ensure the correct spatial orientation of L-Fuc and D-Gal in the bioactive conformation, that is, its role is to pre-organize the Le^x core. D-GlcNAc itself has only weak contacts with the target protein (Figure 2 B).^[13] A comparable role is attributed to D-Neu5Ac, which only contributes to binding through a salt bridge involving its carboxylate group.^[13] Consequently, mimics of D-GlcNAc and D-Neu5Ac were designed to stabilize the bioactive conformation and to keep the entropic costs for binding low or virtually the same as for the highly pre-organized sLe^x (**1**). With (*R,R*)-cyclohexane-1,2-diol (**1**→**3**) which was demonstrated to be a moderate mimic of D-GlcNAc,^[7a] substantial entropy costs arose ($-T\Delta\Delta S$: 3.5 kJ mol⁻¹, Figure 3). When D-Neu5Ac in **3** was replaced by (*S*)-cyclohexyl lactic acid (**3**→**6**, $-T\Delta\Delta S$: 0.6 kJ mol⁻¹), only a small entropy penalty emerged. Finally, (1*R*,2*R*,3*S*)-3-methylcyclohexane-1,2-diol (**6**→**2**) proved to be an optimal replacement of D-GlcNAc (Figure 3),^[7a] resulting in an entropy term similar to that of sLe^x (**1**; Table 1). Compared to antagonist **6**, the improved pre-organization of the core conformation in **2** led to a substantial reduction of the entropy costs ($-T\Delta\Delta S$: -2.4 kJ mol⁻¹). Because the investigated mimics are significantly less polar than D-GlcNAc and D-Neu5Ac, an additional effect needs to be taken into account, namely a substantial alteration of the solvation properties. Although only partial desolvation is necessary, the desolvation of the carbocyclic mimics is enthalpically less unfavorable compared to that of the more polar D-GlcNAc or D-Neu5Ac moieties.^[31]

Strikingly, the introduction of (*R,R*)-cyclohexane-1,2-diol (**1**→**3**) has the same relative effect on enthalpy and entropy as the exchange of D-Neu5Ac for (*S*)-cyclohexyllactic acid (**1**→**5**), which is reflected in the same slope in the entropy–enthalpy plot in Figure 4. In both cases, a significant gain in

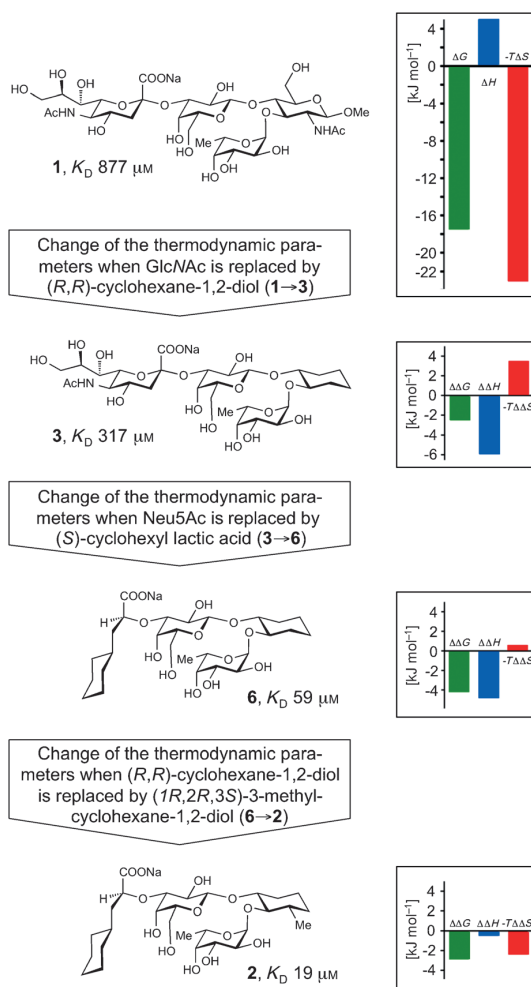


Figure 3. Thermodynamic signature (ΔG , ΔH , $-T\Delta S$) of sLe^x (**1**) (the corresponding data for the antagonists **2–6** are summarized in Table 1) and the changes of the thermodynamic parameters ($\Delta\Delta G$, $\Delta\Delta H$, and $-T\Delta\Delta S$) when D-GlcNAc in **1** is replaced by (*R,R*)-cyclohexane-1,2-diol (**1**→**3**), D-Neu5Ac in antagonist **3** by (*S*)-cyclohexyl lactic acid (**3**→**6**), and (*R,R*)-cyclohexane-1,2-diol in antagonist **6** by (1*R*,2*R*,3*S*)-3-methylcyclohexane-1,2-diol (**6**→**2**).

enthalpy is partially compensated by a loss in entropy. Furthermore, when both mimics are combined in one molecule (**3**→**6**), the effect is not additive, that is, less entropy is lost but also less enthalpy is gained than expected (**6** vs **6**_{expected}, Figure 4). Clearly, the exchange of the carbohydrate moieties does not only change local conformational and solvation properties, but rather the properties of the entire ligand.

In summary, thermodynamic binding parameters for the interaction of E-selectin with sLe^x (**1**) and the glycomimetics **2–6** were investigated by ITC. The interaction of sLe^x with E-selectin is driven by a large favorable entropy term which is partially compensated by an unfavorable enthalpy contribution. The exchange of residues acting as scaffolds with less

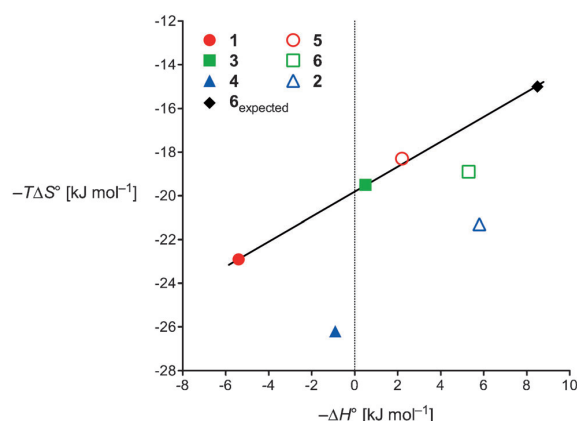


Figure 4. Entropy–enthalpy plot for ligands **1** to **6** and the values expected for **6** (**6_{expected}**) in case the effects caused by the mimetic replacements of D-GlcNAc by cyclohexane-1,2-diol and D-Neu5Ac by (S)-cyclohexyllactic acid were additive.

polar mimics, that is, D-Neu5Ac for (S)-cyclohexyllactic acid and D-GlcNAc for (R,R)-cyclohexane-1,2-diol, resulted in improved binding enthalpy, however accompanied by a loss of binding entropy. Only for mimetic structures that maintain the pre-organization of sLe^x (**1**) in its bioactive conformation, as it is the case for the replacement of D-GlcNAc by (1R,2R,3S)-3-methylcyclohexane-1,2-diol (**1**→**4** or **5**→**2**), a similar entropy term was found. Overall, the almost 50-fold improved affinity of **2** compared to **1** results from a gain in binding enthalpy, whereas the binding entropy is not significantly changed.

The development of glycomimetics with improved binding properties is intrinsically difficult because of the similarity of the ligand (carbohydrate) and the solvent (water). The results of this thermodynamic study suggest, that for a successful development of glycomimetics, carbohydrate moieties with predominantly structural tasks and no or only weak contacts with the target protein should be replaced by hydrophobic mimics, resulting in reduced desolvation penalties and therefore improved enthalpic contributions to binding. In addition, the mimetic replacement should contribute to an improved pre-organization of the binding conformation to optimize the entropy term as well. When the carbohydrate ligand is already almost optimally pre-organized in solution, as it is the case for sLe^x (**1**), the identification of such mimics is a most challenging task.^[7a]

Received: April 2, 2012

Published online: July 2, 2012

Keywords: carbohydrates · E-selectin · glycomimetics · isothermal titration calorimetry · sialyl Lewis^x

- [1] a) B. Ernst, J. L. Magnani, *Nat. Rev. Drug Discovery* **2009**, *8*, 661–677; b) N. Kaila, B. E. Thomas, *Med. Res. Rev.* **2002**, *22*, 566–601; c) E. E. Simanek, G. J. McGarvey, J. A. Jablonowski, C. H. Wong, *Chem. Rev.* **1998**, *98*, 833–862.

- [2] a) C. T. Öberg, H. Leffler, U. J. Nilsson, *Chimia* **2011**, *65*, 18–23; b) L. Ingrassia, I. Camby, F. Lefranc, V. Mathieu, P. Nshimyumukiza, F. Darro, R. Kiss, *Curr. Med. Chem.* **2006**, *13*, 3513–3527.
- [3] P. R. Crocker, J. C. Paulson, A. Varki, *Nat. Rev. Immunol.* **2007**, *7*, 255–266.
- [4] S. A. Mousa, D. A. Cheresh, *Drug Discovery Today* **1997**, *2*, 187–199.
- [5] a) G. S. Kansas, *Blood* **1996**, *88*, 3259–3287; b) D. B. Cines, E. S. Pollak, J. Loscalzo, C. A. Buck, G. A. Zimmerman, R. P. McEver, J. S. Pober, T. M. Wick, B. A. Konkle, B. S. Schwartz, E. S. Barnathan, K. R. McCrae, B. A. Hug, A.-M. Schmidt, D. M. Stern, *Blood* **1998**, *91*, 3527–3561.
- [6] M. L. Phillips, E. Nudelman, F. C. A. Gaeta, M. Perez, A. K. Singhal, S. I. Hakomori, J. C. Paulson, *Science* **1990**, *250*, 1130–1132; G. Walz, A. Aruffo, W. Kolamus, M. Bevilacqua, B. Seed, *Science* **1990**, *250*, 1132–1134.
- [7] a) D. Schwizer, J. Patton, B. Cutting, M. Smiesko, B. Wagner, A. Kato, C. Weckerle, F. P. C. Binder, S. Rabbani, O. Schwardt, J. L. Magnani, B. Ernst, *Chem. Eur. J.* **2012**, *18*, 1342–1351; b) P. W. Bedard, N. Kaila, *Expert Opin. Ther. Pat.* **2010**, *20*, 781–793; c) N. Kaila, B. E. Thomas, *Expert Opin. Ther. Pat.* **2003**, *13*, 305–317; d) G. Thoma, R. Banteli, W. Jahnke, J. L. Magnani, J. T. Patton, *Angew. Chem.* **2001**, *113*, 3756–3759; *Angew. Chem. Int. Ed.* **2001**, *40*, 3644–3647; e) G. Thoma, J. L. Magnani, J. T. Patton, B. Ernst, W. Jahnke, *Angew. Chem.* **2001**, *113*, 1995–1999; *Angew. Chem. Int. Ed.* **2001**, *40*, 1941–1945; f) H. C. Kolb, B. Ernst, *Chem. Eur. J.* **1997**, *3*, 1571–1578.
- [8] a) W. Stahl, U. Sprengard, G. Kretzschmar, H. Kunz, *Angew. Chem.* **1994**, *106*, 2186–2188; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 2096–2098; b) J. Y. Ramphal, Z. L. Zheng, C. Perez, L. E. Walker, S. A. Defrees, F. C. A. Gaeta, *J. Med. Chem.* **1994**, *37*, 3459–3463; c) B. K. Brandley, M. Kiso, S. Abbas, P. Nikrad, O. Srivasatava, C. Foxall, Y. Oda, A. Hasegawa, *Glycobiology* **1993**, *3*, 633–641; d) D. Tyrrell, P. James, N. Rao, C. Foxall, S. Abbas, F. Dasgupta, M. Nashed, A. Hasegawa, M. Kiso, D. Asa, J. Kidd, B. K. Brandley, *Proc. Natl. Acad. Sci. USA* **1991**, *88*, 10372–10376.
- [9] D. V. Erbe, B. A. Wolitzky, L. G. Presta, C. R. Norton, R. J. Ramos, D. K. Burns, J. M. Rumberger, B. N. N. Rao, C. Foxall, B. K. Brandley, L. A. Lasky, *J. Cell Biol.* **1992**, *119*, 215–227.
- [10] a) R. Harris, G. R. Kiddle, R. A. Field, M. J. Milton, B. Ernst, J. L. Magnani, S. W. Homans, *J. Am. Chem. Soc.* **1999**, *121*, 2546–2551; b) L. Poppe, G. S. Brown, J. S. Philo, P. V. Nikrad, B. H. Shah, *J. Am. Chem. Soc.* **1997**, *119*, 1727–1736; c) K. Scheffler, J. R. Brisson, R. Weisemann, J. L. Magnani, W. T. Wong, B. Ernst, T. Peters, *J. Biomol. NMR* **1997**, *9*, 423–436; d) K. Scheffler, B. Ernst, A. Katopodis, J. L. Magnani, W. T. Wang, R. Weisemann, T. Peters, *Angew. Chem.* **1995**, *107*, 2034–2037; *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1841–1844.
- [11] M. Rinnbauer, B. Ernst, B. Wagner, J. Magnani, A. J. Benie, T. Peters, *Glycobiology* **2003**, *13*, 435–443.
- [12] T. Ishida, *J. Phys. Chem. B* **2010**, *114*, 3950–3964.
- [13] a) W. S. Somers, J. Tang, G. D. Shaw, R. T. Camphausen, *Cell* **2000**, *103*, 467–479; b) B. J. Graves, R. L. Crowther, C. Chandran, J. M. Rumberger, S. Li, K.-S. Huang, D. H. Presky, P. C. Familletti, B. A. Wolitzky, D. K. Burns, *Nature* **1994**, *367*, 532–538.
- [14] M. K. Wild, M. C. Huang, U. Schulze-Horsel, P. A. van der Merwe, D. Vestweber, *J. Biol. Chem.* **2001**, *276*, 31602–31612.
- [15] a) J. E. Ladbury, *Biochem. Soc. Trans.* **2010**, *38*, 888–893; b) J. E. Ladbury, G. Klebe, E. Freire, *Nat. Rev. Drug Discovery* **2010**, *9*, 23–27.
- [16] G. A. Holdgate, W. H. Ward, *Drug Discovery Today* **2005**, *10*, 1543–1550.
- [17] a) J. E. DeLorbe, J. H. Clements, M. G. Teresk, A. P. Benfield, H. R. Plake, L. E. Millspaugh, S. F. Martin, *J. Am. Chem. Soc.*

The binding free energy (ΔG) associated with a protein–ligand interaction is composed of enthalpic (ΔH) and entropic ($-T\Delta S$) contributions ($\Delta G = \Delta H - T\Delta S$). The binding energy under standard conditions (ΔG°), where all reactants and products are at a concentration of 1 mol L^{-1} , is calculated from the dissociation constant K_D using the equation $\Delta G = RT \ln K_D$. With isothermal titration calorimetry (ITC),^[16] K_D and the enthalpy ΔH are measured directly if no changes in the protonation states occur during the interaction. The enthalpic term (ΔH) represents the contribution of non-covalent interactions upon binding,^[15b] that is, hydrogen bonds, electrostatic, and dipole–dipole interactions between ligand and receptor.^[17] The entropy term can be dissected into translational and rigid-body rotational entropy,^[18] solvation entropy,^[19] and conformational entropy.^[20]

For our study, an E-selectin/IgG construct consisting of the lectin domain, the EGF-like domain, and six short consensus repeats fused to the Fc part of human IgG1 was used.^[21] The 148 kDa protein was expressed in Chinese Hamster Ovarian cells and purified from the conditioned culture medium by affinity chromatography, first with protein A-Sepharose, followed by a second functional purification with the monoclonal anti-hE-selectin antibody 7A9 (see the Supporting Information). The high degree of purity and functionality of the protein is reflected by the stoichiometry (N) of the calorimetric experiments (Table 1). Batches of up to 50 mg E-selectin/IgG were necessary for ITC measurement to reach c values close to 1.

Our calorimetric investigation had two goals: first, the determination of K_D values of a series of E-selectin ligands and their comparison with data collected by a competitive binding assay (Table 1)^[22] and second, the elucidation of the thermodynamic fingerprints of these ligands. The K_D value

for sLe^x (**1**) binding to E-selectin determined by ITC is (878 ± 93) μM and is thereby in good agreement with previously reported data (e.g. 1.1 to 2.0 mM,^[23] (0.7 ± 0.4) mM^[10b]). In addition, the relative K_D values (rK_D) for the antagonists **2** to **6** also nicely correlate with their relative IC_{50} values ($r\text{IC}_{50}$). In analogy to earlier findings,^[7a] replacement of D-GlcNAc with carbocyclic mimics enhanced binding affinity up to 25-fold (**1**→**4**), whereas the replacement of D-Neu5Ac by (*S*)-cyclohexyllactic acid improved binding 2- to 5-fold (**1**→**5**; **3**→**6**; **4**→**2**).

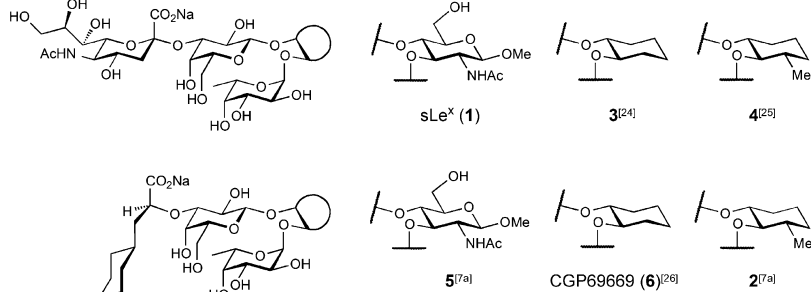
Except for some isolated cases,^[14,27] lectin–oligosaccharide interactions are typically enthalpy driven with mostly unfavorable entropies.^[28] In contrast, the binding of sLe^x to E-selectin is driven by a large entropy term ($-T\Delta S = -23 \text{ kJ mol}^{-1}$). Clearly, the entropy costs caused by the loss of translational and rotational degrees of freedom and conformational changes of ligand and protein upon binding are overcompensated by the beneficial entropy arising from the release of bound water molecules.^[29] This argumentation is supported by two experimental observations. First, the bound conformation was identified as one of two low-energy solution conformations of sLe^x,^[30] demanding only minor conformational adjustments upon binding. Second, the comparison of the crystal structure of apo-E-selectin and E-selectin bound to sLe^x revealed only minor conformational differences.^[13]

The beneficial entropy term, however, is partially compensated by an unfavorable change in enthalpy. To enable the pharmacophoric groups of sLe^x (**1**) to interact with their target, a predominantly polar surface area of approximately 275 \AA^2 ^[13a] on both interacting moieties has to be desolvated. Because the newly formed polar interactions between the pharmacophores of sLe^x and E-selectin do not fully compensate for the desolvation penalty of the polar binding interface (Figure 2),^[31] a net loss of enthalpy ($\Delta H^\circ = +5.4 \text{ kJ mol}^{-1}$) is observed. Thus, the directed polar interactions of the pharmacophores contribute to specificity rather than affinity.

Thus, sLe^x (**1**) represents a surrogate of clustered water molecules attached to a scaffold. As a “pre-organized water oligomer” it offers an array of directed hydrogen bonds for the highly specific binding to E-selectin. The clear entropic benefit of the release of water molecules from the large binding interface to bulk water and the high degree of pre-organization of sLe^x result in the observed large entropy gain which provides the impetus for the binding process.

The concept of conformational pre-organization was also exploited for the development of selectin antagonists (Figure 3). More precisely, in sLe^x (**1**), D-GlcNAc acts as

Table 1: Affinity and thermodynamic parameters for the interaction of **1**–**6** with E-selectin.^[a]



Ligand	$r\text{IC}_{50}$	rK_D	K_D [μM]	ΔG [kJ mol^{-1}]	ΔH [kJ mol^{-1}]	$-T\Delta S$ [kJ mol^{-1}]	N
1	1	1	878 ± 93	-17.5 ± 0.2	$+5.4 \pm 0.7$	-23 ± 1	1
3	0.3	0.36	317	-20.0	-0.5	-19.5	1
4	0.05	0.04	38	-25.3	$+0.9$	-26.2	0.94
5	0.27	0.30	260	-20.5	-2.2	-18.3	1
6	0.08	0.07	59 ± 4	-24.2 ± 0.2	-5.3 ± 0.4	-18.9 ± 0.6	0.93 ± 0.08
2	0.014	0.02	19 ± 2	-27.1 ± 0.2	-5.8 ± 0.1	-21.3 ± 0.4	0.97 ± 0.01

[a] Relative IC_{50} values ($r\text{IC}_{50}$) and relative K_D values (rK_D) are reported relative to the reference compound sLe^x (**1**). IC_{50} values were determined in a competitive binding assay.^[22] K_D and ΔH were measured in ITC experiments, ΔG , and $T\Delta S$ were calculated according to the equations $\Delta G = \Delta H - T\Delta S$ and $\Delta G = RT \ln K_D$. N = stoichiometric ratio of ligand and protein.

- 2009, 131, 16758–16770; b) M. C. Chervenak, E. J. Toone, *J. Am. Chem. Soc.* **1994**, 116, 10533–10539.
- [18] A. V. Finkelstein, J. Janin, *Protein Eng.* **1989**, 3, 1–3.
- [19] a) K. P. Murphy, *Biophys. Chem.* **1994**, 51, 311–326; b) R. L. Baldwin, *Proc. Natl. Acad. Sci. USA* **1986**, 83, 8069–8072.
- [20] K. K. Frederick, M. S. Marlow, K. G. Valentine, A. J. Wand, *Nature* **2007**, 448, 325–330.
- [21] W. Jahnke, H. C. Kolb, M. J. J. Blommers, J. L. Magnani, B. Ernst, *Angew. Chem.* **1997**, 109, 2715–2719; *Angew. Chem. Int. Ed. Engl.* **1997**, 36, 2603–2607.
- [22] a) G. Weitz-Schmidt, G. Stokmaier, G. Scheel, N. E. Nifant'ev, B. Tuzikov, N. V. Bovin, *Anal. Biochem.* **1996**, 238, 184–190; b) G. Thoma, J. L. Magnani, R. Oehrlein, B. Ernst, F. Schwarzenbach, R. O. Duthaler, *J. Am. Chem. Soc.* **1997**, 119, 7414–7415.
- [23] R. M. Cooke, R. S. Hale, S. G. Lister, G. Shah, M. P. Weir, *Biochemistry* **1994**, 33, 10591–10596.
- [24] A. Toepfer, G. Kretschmar, E. Bartnik, *Tetrahedron Lett.* **1995**, 36, 9161–9164.
- [25] For the synthesis of **4**, see the Supporting Information.
- [26] K. E. Norman, G. P. Anderson, H. C. Kolb, K. Ley, B. Ernst, *Blood* **1998**, 91, 475–483.
- [27] a) C. O. Sallum, R. A. Kammerer, A. T. Alexandrescu, *Biochemistry* **2007**, 46, 9541–9550; b) M. Kapoor, H. Srinivas, E. Kandiah, E. Gemma, L. Ellgaard, S. Oscarson, A. Helenius, A. Surolia, *J. Biol. Chem.* **2003**, 278, 6194–6200; c) P. G. Rani, K. Bachhawat, G. B. Reddy, S. Oscarson, A. Surolia, *Biochemistry* **2000**, 39, 10755–10760.
- [28] a) M. Ambrosi, N. R. Cameron, B. G. Davis, *Org. Biomol. Chem.* **2005**, 3, 1593–1608; b) T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, 102, 387–429; c) T. K. Dam, B. S. Cavada, T. B. Grangeiro, C. F. Santos, V. M. Ceccatto, F. A. M. de Sousa, S. Oscarson, C. F. Brewer, *J. Biol. Chem.* **2000**, 275, 16119–16126; d) E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, 4, 719–728.
- [29] a) P. R. Connelly in *Structure-Based Drug Design: Thermodynamics, Modeling, and Strategy* (Eds.: J. E. Ladbury, P. R. Connelly), Springer, Berlin, **1997**, pp. 143–157; b) J. D. Dunitz, *Science* **1994**, 264, 670.
- [30] a) T. J. Rutherford, D. G. Spackman, P. J. Simpson, S. W. Homans, *Glycobiology* **1994**, 4, 59–68; b) Y. Ichikawa, Y. C. Lin, D. P. Dumas, G. J. Shen, E. Garciajunceda, M. A. Williams, R. Bayer, C. Ketcham, L. E. Walker, J. C. Paulson, C. H. Wong, *J. Am. Chem. Soc.* **1992**, 114, 9283–9298; c) Y. C. Lin, C. W. Hummel, D. H. Huang, Y. Ichikawa, K. C. Nicolaou, C. H. Wong, *J. Am. Chem. Soc.* **1992**, 114, 5452–5454.
- [31] a) A. J. Ruben, Y. Kiso, E. Freire, *Chem. Biol. Drug Des.* **2006**, 67, 2–4; b) P. R. Connelly, R. A. Aldape, F. J. Bruzzese, S. P. Chambers, M. J. Fitzgibbon, M. A. Fleming, S. Itoh, D. J. Livingston, M. A. Navia, J. A. Thomson, K. P. Wilson, *Proc. Natl. Acad. Sci. USA* **1994**, 91, 1964–1968; c) S. Cabani, P. Gianni, V. Mollica, L. Lepori, *J. Solution Chem.* **1981**, 10, 563–595.
- [32] Maestro, Version 9.1, Schrödinger, LLC, New York, NY, **2010**.



Supporting Information

© Wiley-VCH 2012

69451 Weinheim, Germany

Sialyl Lewis^x: A “Pre-Organized Water Oligomer”?**

*Florian P. C. Binder, Katrin Lemme, Roland C. Preston, and Beat Ernst**

anie_201202555_sm_miscellaneous_information.pdf

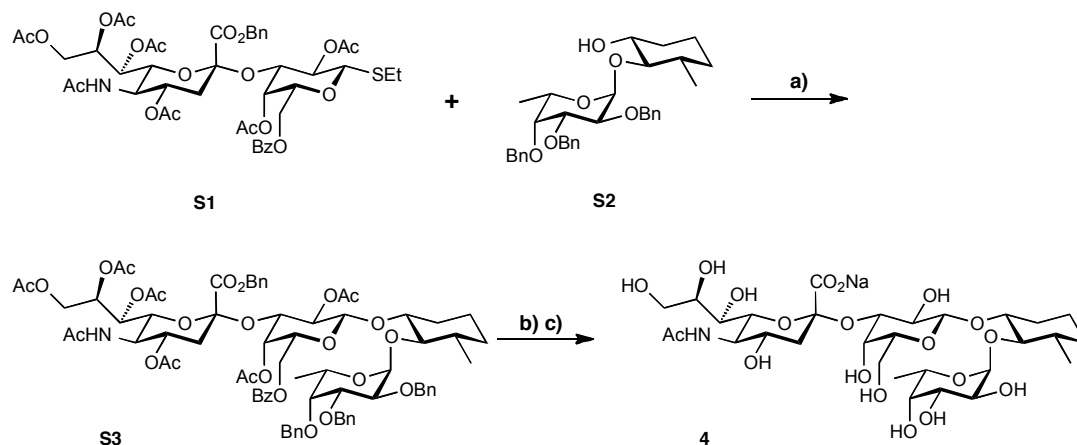
Contents

1. Synthesis of ligand 4	2
2. Materials and methods	5
3. Generation of the 7A9 antibody column	5
4. E-selectin/IgG expression and purification	6
5. Competitive binding assay	7
6. Isothermal titration calorimetry	7
12. Graphics	8
12. References	12
13. Experimental data	13

1. Synthesis of ligand 4

General Information

NMR spectra were recorded on a Bruker Avance DMX-500 (500 MHz) spectrometer. Assignment of ^1H and ^{13}C NMR spectra was achieved using 2D methods (COSY, HSQC, HMQC, HMBC). Chemical shifts are given in ppm and were assigned in relation to the solvent signals on the d-scale^[S1] or to tetramethylsilane (0 ppm) as internal standard. Coupling constants J are given in Hertz (Hz). Multiplicities were specified as follows: s (singlet), d (doublet), dd (doublet of a doublet), t (triplet), q (quartet), m (multiplet). For assignment of resonance signals to the appropriate nuclei the following abbreviations were used: Fuc (fucose), Gal (galactose), MeCy (3-methylcyclohexane-1,2-diol), Sia (sialic acid). Reactions were monitored by TLC using glass plates coated with silica gel 60 F₂₅₄ (Merck) and visualized by using UV light and/or by charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10% H₂SO₄). Column chromatography was performed using the RediSep Companion from Teledyne Isco with normal phase RediSep columns from the same manufacturer or reversed-phase columns containing LiChroprep RP-18 (40-63 μm) from Merck KGaA, Darmstadt, Germany. Size exclusion chromatography was performed with Bio-Gel[®] P-2 Gel (45-90 mm) from Bio-Rad. Solvents were purchased from Sigma-Aldrich. Dichloromethane (CH₂Cl₂) was dried by filtration over Al₂O₃ (Fluka, type 5016 A basic). Optical rotations were measured using a Perkin-Elmer Polarimeter 341. Electron spray ionization mass spectra (ESI-MS) were obtained on a Waters micromass ZQ. HRMS analysis were carried out using a Agilent 1100 LC equipped with a photodiode array detector and a Micromass QTOF I equipped with a 4 GHz digital-time converter. Microanalysis was performed at the Institute of Organic Chemistry at the University of Basel, Switzerland. Purity of final compounds was determined on an Agilent 1100 HPLC; detector: ELS, Waters 2420; column: Waters Atlantis dC18, 3 μm , 4.6 x 75 mm; eluents: A: water + 0.1% TFA; B: 90% acetonitrile + 10% water + 0.1% TFA; linear gradient: 0 - 1 min 5% B; 1 - 20 min 5 to 70% B; flow: 0.5 mL/min.



Scheme a) DMTST, CH₂Cl₂, MS 4 Å, RT, 3 d, 55%; b) H₂, Pd(OH)₂/C, dioxane, H₂O, RT, 12 h; c) aq. NaOH, MeOH, RT, 81% from **S3**.

Donor **S1**^[S2] was coupled to pseudodisaccharide **S2**^[S3] using dimethyl(methylthio)sulfonium triflate (DMTST)^[S4] as promotor (Scheme). Hydrogenolytic debenzylation followed by saponification with sodium hydroxide afforded **4**.

(Benzyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero- α -D-galacto-2-nonulopyranosynate)-(2 \rightarrow 3)-2,4-di-*O*-acetyl-6-*O*-benzoyl- α -D-galactopyranosyl-(1 \rightarrow 1)-[2,3,4-tri-*O*-benzyl- α -L-fucopyranosyl-(1 \rightarrow 2)]-(1*R*,2*R*,3*S*)-3-methyl-cyclohexane-1,2-diol (S3**).**

Compound **S1** (300 mg, 0.31 mmol) and **S2** (256 mg, 0.47 mmol) were dissolved in anhydrous CH₂Cl₂ (6.0 mL). Powdered activated molecular sieves 4 Å (0.6 g) were added and the mixture was stirred at RT under argon. After 3.5 h, a solution of DMTST (246 mg, 0.95 mmol) in anhydrous CH₂Cl₂ (1.9 mL) that had been stirred with molecular sieves 4 Å (0.19 g) for 3.5 h, was added. After stirring for 3 d, the solution was diluted with CH₂Cl₂ (40 mL), filtered, and successively washed with satd. aq. NaHCO₃ (50 mL) and brine (50 mL). The aqueous layers were extracted with CH₂Cl₂ (2 \times 50 mL) and the combined organic layers were dried over Na₂SO₄ and concentrated under reduced pressure. Column chromatography on silica (PE/EtOAc/MeOH 8/5/0.5 to 8/5/0.7) afforded **S3** as white foam (248 mg, 0.17 mmol, 55%).

$[\alpha]_D^{20}$ -18.9° (c 1.01, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ 8.04-7.14 (m, 25H, Ar-H), 5.48 (ddd, ³*J* = 2.6, 6.2, 9.1 Hz, 1H, Sia-H8), 5.21 (A of AB, ²*J* = 12.0 Hz, 1H, PhCH₂), 5.20 (dd, ³*J* = 2.8, 9.4 Hz, 1H, Sia-H7), 5.01 (d, ³*J* = 3.7 Hz, 2H, Fuc-H1, Gal-H4), 4.90 (A' of A'B', *J* = 12.0 Hz, 1H, PhCH₂), 4.88-4.64 (m, 10H, Fuc-H5, Gal-H2, SiaNH, Sia-H4, 3

PhCH₂), 4.56 (d, ³*J* = 8.0 Hz, 1H, Gal-H1), 4.48 (dd, ³*J* = 3.3, 10.2 Hz, 1H, Gal-H3), 4.25 (dd, ³*J* = 2.6 Hz, ²*J* = 12.4 Hz, 1H, Sia-H9a), 4.17 (dd, ³*J* = 6.4 Hz, ²*J* = 10.8 Hz, 1H, Gal-H6a), 4.05 (dd, ³*J* = 3.7, 10.3 Hz, 2H, Fuc-H2), 4.02–3.93 (m, 3H, Fuc-H3, Gal-H6b, Sia-H5), 3.90–3.83 (m, 2H, Gal-H5, Sia-H9b), 3.59 (m, 1H, Fuc-H4), 3.50 (m, 1H, MeCy-H1), 3.37 (dd, 1H, ³*J* = 2.7, 10.8 Hz, Sia-H6), 3.18 (t, ³*J* = 9.2 Hz, MeCy-H2), 2.50 (dd, ³*J* = 4.6 Hz, ²*J* = 12.7 Hz, 1H, Sia-H3^{eq}), 2.09, 2.05 (2s, 6H, 2 COCH₃), 1.99 (m, 1H, MeCy), 1.99, 1.94, 1.90, 1.74, 1.72 (5s, 15H, 5 COCH₃), 1.65–1.47 (m, 4H, Sia-H3^{ax}, MeCy), 1.21–1.11 (m, 5H, Fuc-H6, MeCy), 1.03 (d, ³*J* = 6.4 Hz, 3H, MeCy-CH₃), 0.98 (m, 1H, MeCy); ¹³C NMR (125.8 MHz, CDCl₃): δ 170.8, 170.7, 170.6, 170.4, 169.9, 169.5 (7C, COCH₃), 167.5 (Sia-C1), 165.8 (ArCO), 139.3, 139.1, 138.7, 134.9, 133.3, 128.9, 128.7, 128.6, 128.5, 128.4, 128.3, 127.6, 127.4, 127.3 (30C, Ar-C), 99.4 (Gal-C1), 98.3 (Fuc-C1), 96.9 (Sia-C2), 82.2 (MeCy-C2), 80.6, 80.5 (Fuc-C3, MeCy-C1), 77.8 (Fuc-C4), 76.7 (Fuc-C2), 74.6, 74.3, 72.7 (3C, PhCH₂), 72.0, 71.9 (Gal-C3, Sia-C6), 70.0, 69.8 (Gal-C2, Gal-C5), 69.4 (Sia-C4), 68.4 (PhCH₂), 67.8 (2C, Gal-C4, Sia-C8), 67.1 (Sia C7), 66.3 (Fuc-C5), 62.8 (Sia-C9), 61.5 (Gal-C6), 49.0 (Sia-C5), 39.2 (MeCy-C3), 37.5 (Sia-C3), 33.6 (MeCy-C4), 30.9 (MeCy-C6), 23.3 (CH₃CO), 23.1 (MeCy-C5), 21.5, 21.0, 20.9, 20.8 (6C, COCH₃), 18.9 (MeCy-CH₃), 17.1 (Fuc-C6); MS (ESI) *m/z*: calcd. for C₇₇H₉₁NNaO₂₆ [M+Na]⁺: 1468.6; found: 1468.6; elemental analysis calcd (%) for C₇₇H₉₁NO₂₆ + 0.5 H₂O (1455.55): C 63.54, H 6.37, N 0.96; found: C 63.58, H 6.35, N 0.80.

(Sodium 5-acetamido-3,5-dideoxy-D-glycero-(α -D-galacto-2-nonulopyranosynate)-(2? 3)- α -D-galactopyranosyl-(1? 1)-[α -L-fucopyranosyl-(1? 2)]-(1R,2R,3S)-3-methylcyclohexane-1,2-diol (4).

Compound **S3** (310 mg, 0.21 mmol) was dissolved in dioxane/water (4/1, 10 mL) under argon. Pd(OH)₂/C (40 mg, 10% Pd(OH)₂) was added and the resulting mixture was hydrogenated (4 bar H₂) at RT. After 24 h, the mixture was filtered and the solvent removed under reduced pressure yielding 220 mg of a white solid, which was directly used for saponification. The crude product (75 mg) was stirred in aqueous NaOH (1 N, 1.5 mL) for 24 h at RT, lyophilized, and purified *via* SEC and RP chromatography (H₂O/MeOH). Lyophilization from water afforded **4** as white fluffy solid (42 mg, 0.056 mmol, 81%).

[α]_D²⁰ -47.4° (c 0.89, MeOH); ¹H NMR (500.1 MHz, D₂O): δ 5.07 (d, ³*J* = 3.6 Hz, 1H, Fuc-H1), 5.05–4.71 (m, Fuc-H5), 4.53 (d, ³*J* = 7.8 Hz, 1H, Gal-H1), 4.05 (dd, ³*J* = 2.4, 9.6 Hz, 1H, Gal-H3), 3.92 (m, Gal-H4), 3.90–3.50 (m, 15H, Fuc-H2, Fuc-H3, Fuc-H4, Gal-H2, Gal-H5, Gal-H6a, Gal-H6b, MeCy-H1, Sia-H4, Sia-H5, Sia-H6, Sia-H7, Sia-H8, Sia-H9a, Sia-H9b), 3.20 (t, ³*J* = 9.6 Hz, 1H, MeCy-H2), 2.73 (dd, ³*J* = 4.4 Hz, ²*J* = 12.1 Hz, 1H, Sia-H3^{eq}), 2.13

(m, 1H, MeCy-H6a), 2.00 (s, 3H, COCH₃), 1.78 (t, $^3J = ^2J = 12.1\text{Hz}$, 1H, Sia-H3^{ax}) 1.69-1.52 (m, 3H, MeCy-H3, MeCy-H4a, MeCy-H5a), 1.33-1.17 (m, 2H, MeCy-H5b, MeCy-H6b), 1.15 (d, $^3J = 6.4\text{Hz}$, 1H, Fuc-H6), 1.11-0.99 (m, 4H, MeCy-CH₃, MeCy-H4b); ¹³C NMR (125.8 MHz, D₂O, CD₃OD_{ref}): δ 176.0 (COCH₃), 175.0 (Sia-C1), 100.8 (Sia-C2), 100.5 (Gal-C1), 99.8 (Fuc-C1), 85.0 (MeCy-C2), 79.4 (MeCy-C1), 76.9 (Gal-C3), 75.5 (Gal-C5), 73.8 (Sia-C6), 73.0 (Fuc-C4), 72.7 (Sia-C8), 70.2 (Fuc-C3), 69.9 (Gal-C2), 69.4, 69.2, 69.1 (3C, Fuc-C2, Sia-C4, Sia-C7), 68.5 (Gal-C4), 67.5 (Fuc-C5), 63.6 (Sia-C9), 62.6 (Gal-C6), 52.7 (Sia-C5), 40.6 (Sia-C3), 39.8 (MeCy-C3), 34.2 (MeCy-C4), 31.1 (MeCy-C6), 23.6 (MeCy-C5), 23.0 (COCH₃), 19.2 (MeCy-CH₃), 16.4 (Fuc-C6); HR-MS (ESI) *m/z*: calcd for C₃₀H₅₀NNa₂O₁₉ [M+Na]⁺: 774.2767; found: 774.2768; HPLC-purity ≥ 99.5 %.

2. Materials and methods

Ham'sF-12 medium, RPMI 1640 medium, geneticin (G418) sulfate, αDMEM-medium, peroxidase ABTS (2,2'-azino-bis[3-ethylbenzthiazoline-6-sulfonic acid]) single solution, and FCS (fetal calf serum) were all purchased from Invitrogen (Paisley, UK). Sodium pyruvate, penicillin/streptomycin, HEPES, NaOH, CaCl₂ x 2 H₂O, Tween 20, ammonium acetate, acetic acid, and cyanogen bromide-activated-Sepharose[®] 4B were purchased from Sigma-Aldrich Chemie GmbH (Steinheim, Germany). NaCl and HCl were purchased from Merck KGaA (Darmstadt, Germany). Tris was purchased from AppliChem GmbH (Darmstadt, Germany). Protein A-Sepharose[®] was purchased from BioVision (Mountain View, CA), sLe^a-PAA-biotin was purchased from GlycoTech Corporation (Gaithersburg). The streptavidin peroxidase conjugate was purchased from Roche Diagnostic (Rotkreuz, Switzerland). The BioLogic Duo Flow fast protein liquid chromatography (FPLC) system and the UNO[™] Q ion exchange column were purchased from BioRad (Reinach, Switzerland). The Vivaspin[®] 20 centrifugal concentration tubes were purchased from Sartorius-Stedim (Göttingen, Germany). The SpectraMax 190 microplate reader was obtained from Molecular Devices (Sunnyvale, CA, USA). The VP-ITC apparatus was purchased from MicroCal Inc. (Uppsala, Sweden). Slide-A-Lyzer cassettes were obtained from Thermo Fisher Scientific (Rockford, IL, USA).

3. Generation of the 7A9 antibody column

For the purification of E-selectin/IgG, an affinity column with the functional monoclonal anti-hE-selectin antibody 7A9 was prepared. Mouse hybridoma cells expressing the 7A9 antibody (ATCC No. HB-10135[™]) were cultivated in suspension at 37 °C and 5% CO₂ with

RPMI 1640 culture medium supplemented with 10-15% FCS, 100 U/mL penicillin, and 100 $\mu\text{g/mL}$ streptomycin. The cells were maintained at 10^5 - 10^6 cells/mL at all times. Culture medium was harvested once or twice weekly by centrifugation at 1000 rpm and 4 °C for 10 minutes. The 7A9 antibody was purified using fast protein liquid chromatography, first on a Protein A-sepharose[®] column with the buffers A (50 mM Tris, 150 mM NaCl, pH 7.4, 0.05% Tween 20) for loading, B (5 mM ammonium acetate, pH 5) for washing, C (500 mM acetate, pH 3.0) for elution, and D (2 M Tris) for neutralizing. Higher purity was achieved by subsequent anion exchange chromatography on a Uno[™] Q column with buffers E (20 mM Tris, pH 8.0) and F (20 mM Tris, 1 M NaCl, pH 8.0). Approximately 5 mg of purified antibody was coupled to 1 mL of swollen cyanogen bromide-activated-Sepharose[®] 4B according to the manufacturer's protocol.

4. E-selectin/IgG expression and purification

CHO-K1 cells expressing the E-selectin/IgG construct, which includes the lectin domain, the EGF-like domain, and six complement repeat domains of human E-selectin fused to the Fc part of human IgG1 were generated as previously described.^[S5] The cells were cultivated as monolayers at 37 °C, 5% CO₂, in F-12 culture medium supplemented with 10 mM sodium pyruvate, 10% FCS, 100 U/mL penicillin, 100 $\mu\text{g/mL}$ streptomycin, and 0.4 mg/mL geneticin (G418) sulfate. For the expression, the cells were adapted to 5% FCS and α DMEM-medium with the same additives as for the F-12 culture medium. Conditioned culture medium was harvested once a week and stored at -20 °C.

E-selectin/IgG was purified from the conditioned culture medium by affinity chromatography, first with protein A-Sepharose[®], followed by a second purification with the 7A9-Sepharose[®] column. In both cases, buffers A, B, C, and D were used (see above). Purified E-selectin/IgG was concentrated by ultrafiltration (1610 x g, Vivaspin 20, 50 kDa cut off) and dialysed over night against assay buffer (10 mM HEPES, 150 mM NaCl, 1 mM CaCl₂, pH 7.4) using Slide-A-Lyzer dialysis cassettes (cut off 10 kDa). The protein purity was confirmed by standard SDS-PAGE analysis and the concentration was determined by HPLC-UV against a BSA standard.^[S6]

The binding activity was verified with a carbohydrate polymer-binding assay for the EC₅₀ determination, similar to the protocols described in Ref^[S7] Peroxidase coupled sLe^a-polymer was prepared as follows: 20 μL biotinylated sLe^a (sLe^a-PAA-biotin, 1 mg/mL), 80 μL streptavidin peroxidase conjugate (500 U/mL), 20 μL FCS, and 80 μL assay buffer were combined and incubated for 2 h at 37 °C. The complex is stable for several weeks at 4 °C.

Nunc MaxiSorp™ 96 well plates were coated with E-selectin/IgG (1 $\mu\text{g/ml}$, 100 $\mu\text{L/well}$) over night at 4 °C and were subsequently blocked with BSA (3% w/v in assay buffer, 200 $\mu\text{L/well}$) for 2 h at 4 °C, followed by incubation for 3 h at room temperature with a serial dilution of peroxidase coupled sLe^a-polymer. The excess of sLe^a-polymer was washed, ABTS solution (100 $\mu\text{L/well}$) was added, and the colorimetric reaction was stopped after 10 min with 100 $\mu\text{L/well}$ oxalic acid (2% w/v). The OD was measured at 415 nm with a microplate reader .

5. Competitive binding assay

The competitive binding assay was adapted from Ref^[S71] Most steps were performed according to the EC₅₀ determination (see above). In the incubation step, a constant concentration of sLe^a-polymer (100 ng/mL, 50 $\mu\text{L/well}$), and a serial dilution of antagonist (50 $\mu\text{L/well}$) was applied.

6. Isothermal titration calorimetry

ITC experiments were performed with a VP-ITC instrument. Protein samples were dialyzed over night against assay buffer using Slide-A-Lyzer dialysis cassettes (10 kDa cut-off). Protein and ligand samples were degassed with vacuum prior to the assays. All measurements were performed at 25 °C. Injections of 5 to 15 μL ligand solutions (1.4 - 12 mM) were added from a computer controlled 300 μl microsyringe at an interval of 5 min into the sample cell solution containing E-selectin/IgG (45 to 200 μM , sample cell volume 1.4037 mL or 1.4523 mL) with stirring at 307 rpm. The quantity $c = \text{Mt}(0) K_D^{-1}$, where $\text{Mt}(0)$ is the initial macromolecule concentration, is of importance in titration microcalorimetry and should be between 1 and 1000.^[S8] The c -values for compound **2**, **4**, and **6** were between 1 and 5. The experiments with methyl sLe^x (**1**), **3**, and **5** were performed with c values below 1. In these cases, the stoichiometry N was fixed to 1 to allow for reliable determination of the dissociation constant K_D and the change in enthalpy ΔH .^[S9] Control experiments, i.e. ligand injection into buffer (10 mM HEPES, 150 mM NaCl, 1 mM CaCl₂, pH 7.4) in absence of protein, showed that the heats of dilution were small and constant. The first injection was always excluded from data analysis as it usually suffers from sample loss during the mounting of the syringe and the equilibration preceding the actual titration. Raw data was collected and the area under each peak was integrated, followed by correction for heats of dilution and mixing by subtracting the final baseline consisting of small peaks of the same size to zero. The data were analyzed with ORIGIN Software (Microcal Inc.) by three-parameter data fitting of a single-site binding isotherm, which yields ΔH (enthalpy of

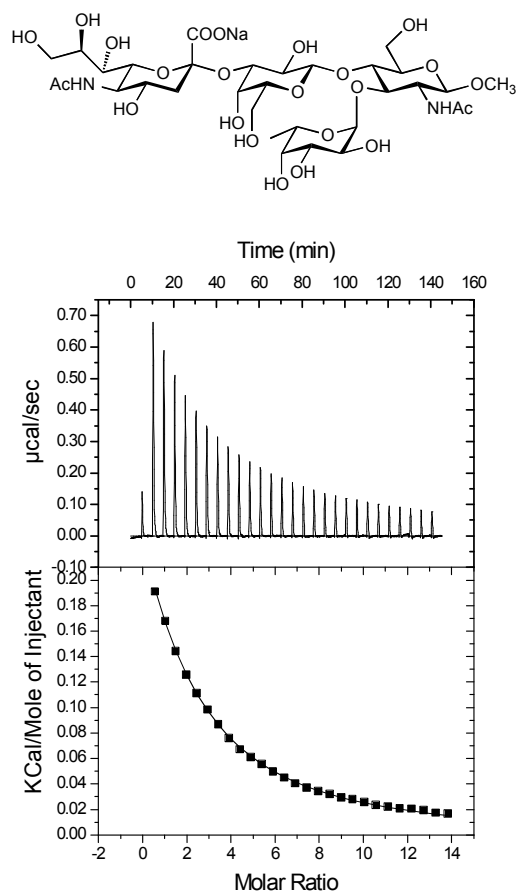
binding), K_D (dissociation constant) and N (stoichiometry) for compounds **2**, **4** and **6**. A two-parameter data fitting of a single-site binding isotherm was used for methyl sLe^x (**1**), compound **3** and **5**, which yields ΔH and K_D .

Thermodynamic parameters were calculated from equation S1.

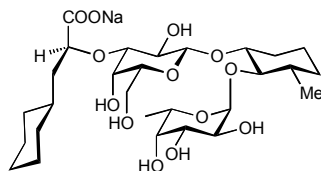
$$\Delta G = \Delta H - T\Delta S = RT \ln K_D \quad (\text{S1})$$

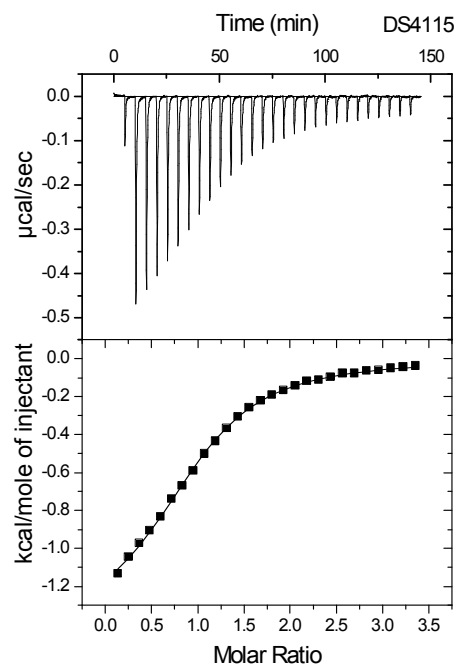
where ΔG , ΔH , and ΔS are the changes in free energy, enthalpy, and entropy, respectively. T is the absolute temperature, and R is the gas constant (8.314 J/molK).

methyl sLe^x (**1**)

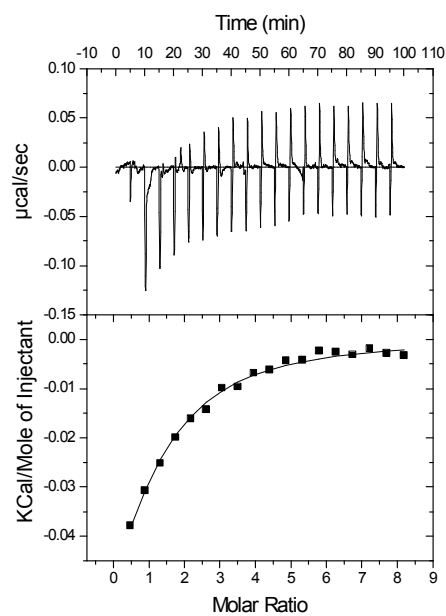
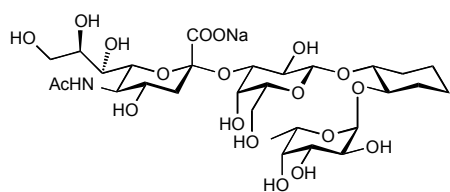


Compound 2

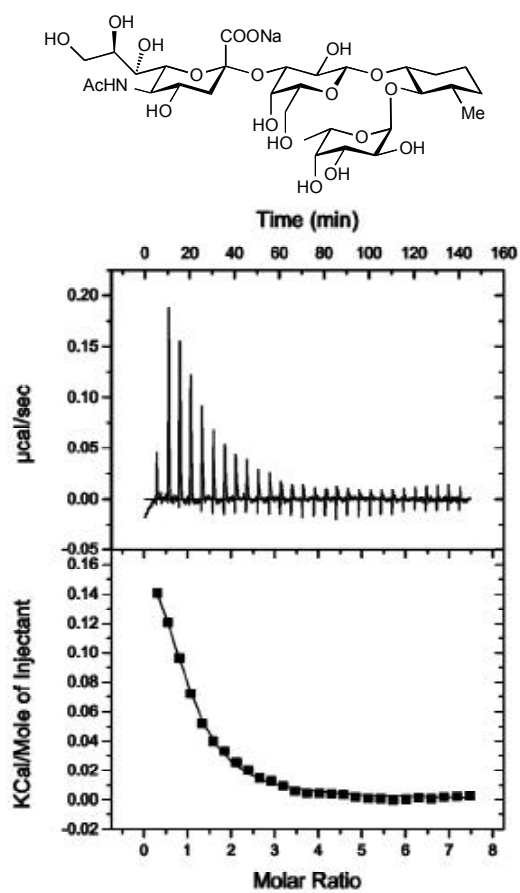
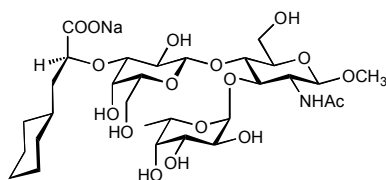


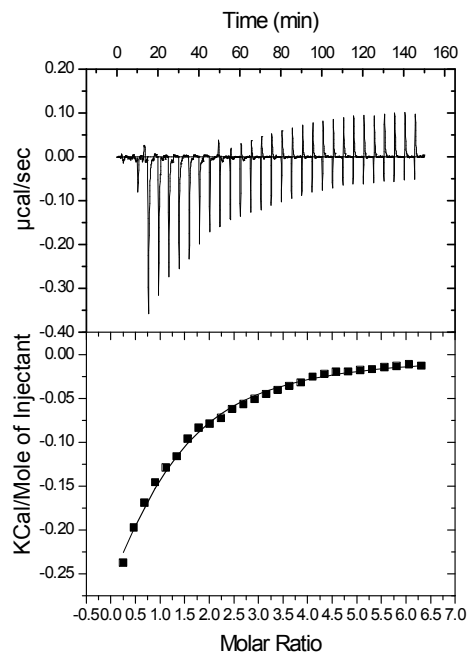


Compound 3

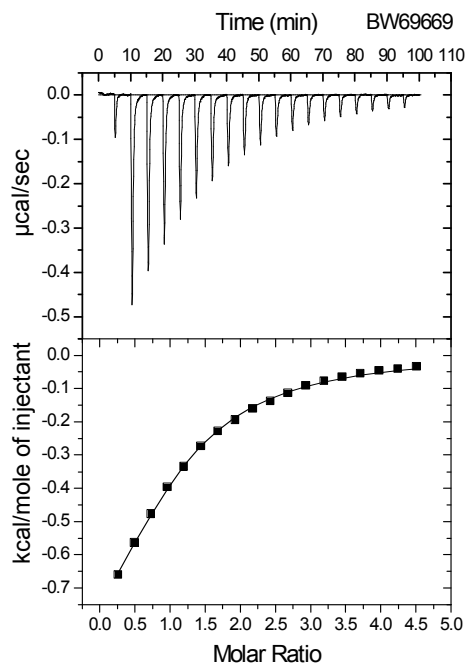
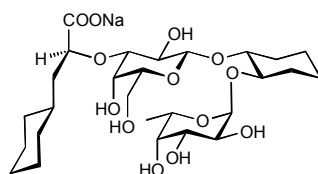


Compound 4

**Compound 5**



Compound 6

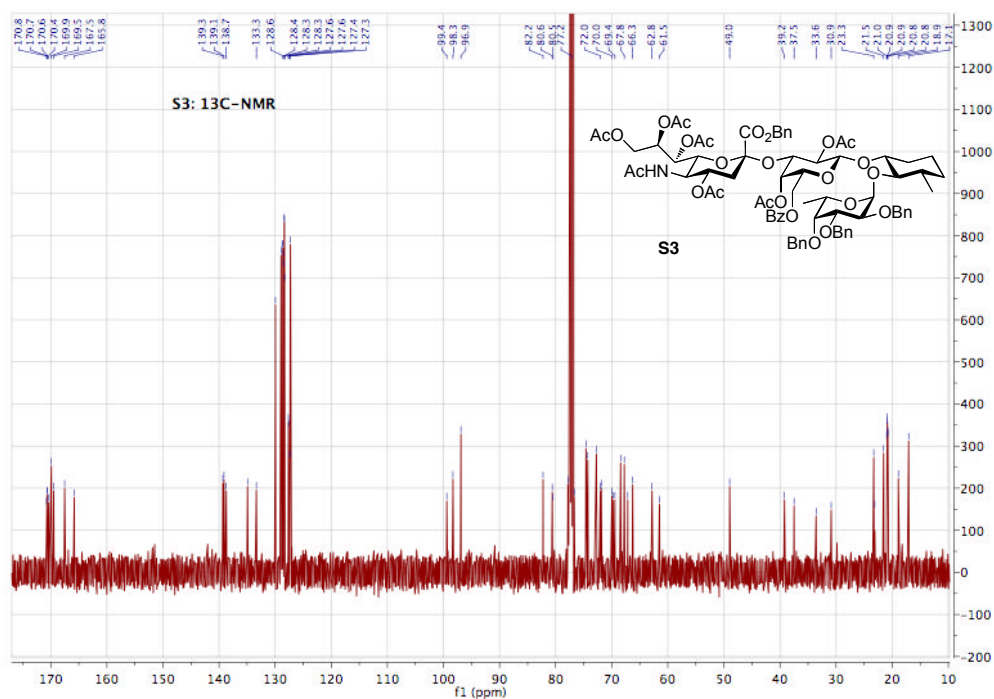
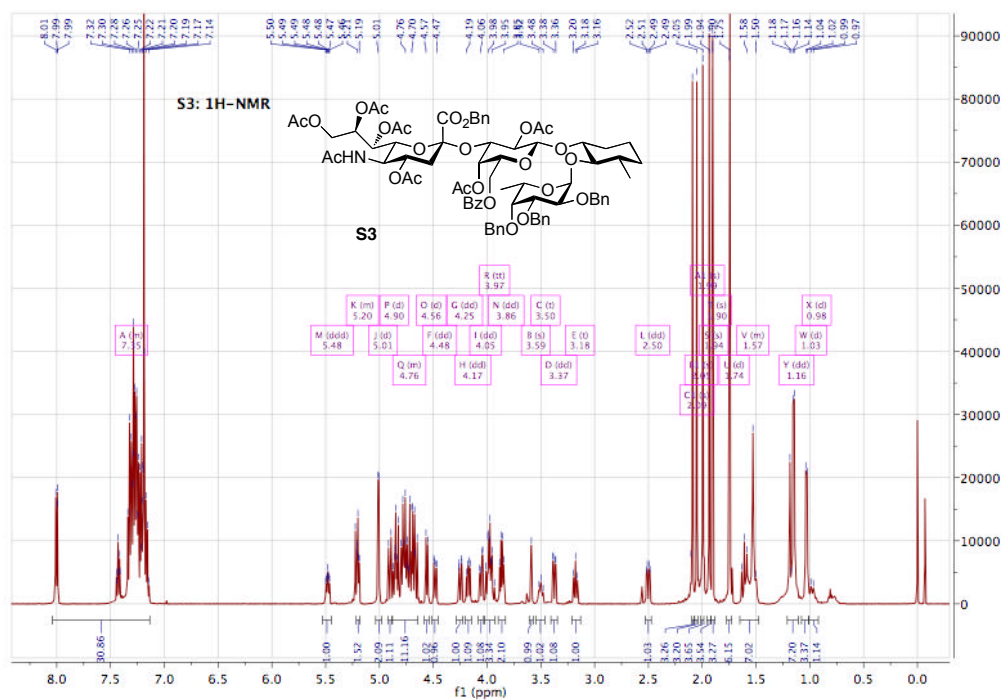


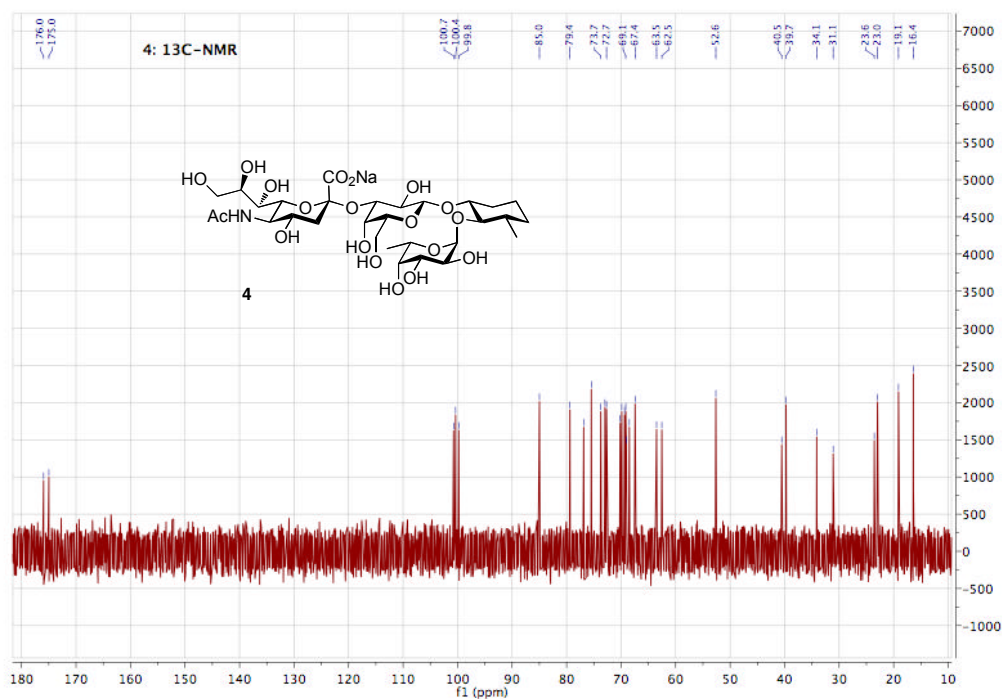
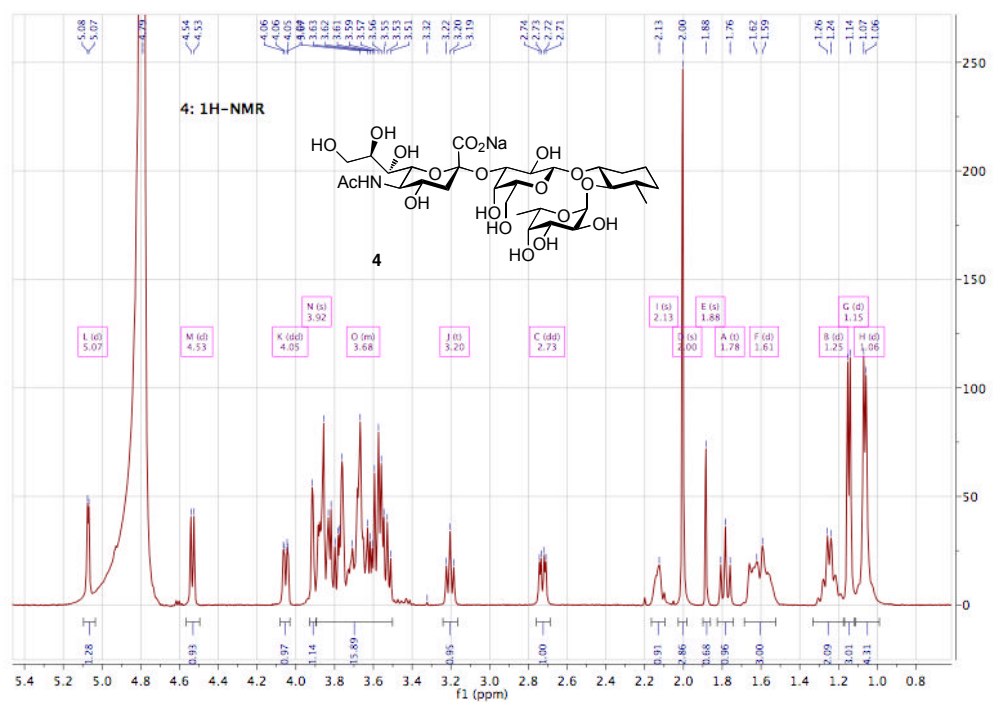
7. Graphics

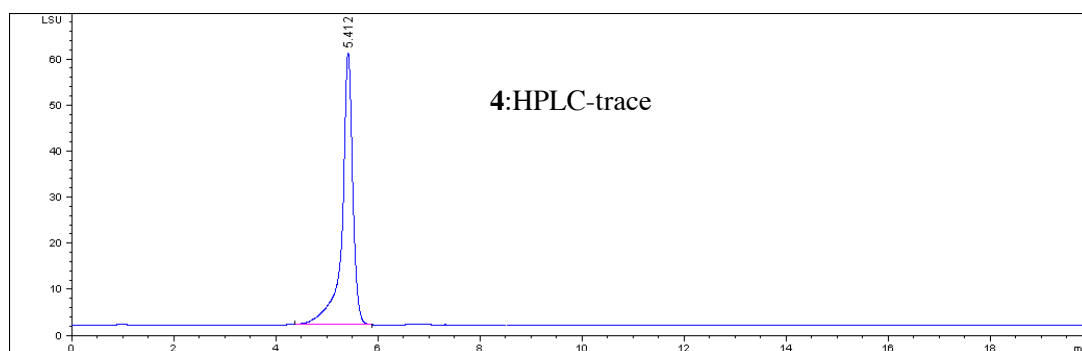
The graphical abstract and Figure 2 are based on the crystal structure of sLe^x in complex with E-selectin (pdb-code 1G1T)^[S10] and illustrated with the software Maestro version 9.1, Schrödinger, LLC, New York, NY, 2010.

8. References

- [S1] H. E. Gottlieb, V. Kotlyar, A. Nudelman, *J. Org. Chem.* **1997**, 62, 7512-7515.
- [S2] A. Bhunia, O. Schwardt, H. Gathje, G. P. Gao, S. Kelm, A. J. Benie, M. Hricovini, T. Peters, B. Ernst, *Chembiochem* **2008**, 9, 2941-2945.
- [S3] D. Schwizer; J. Patton; B. Cutting; M. Smiesko; B. Wagner; A. Kato; C. Weckerle, F.P.C. Binder, S. Rabbani, O. Schwardt, J.L. Magnani, B. Ernst, *Chem. Eur. J.* **2012**, 18, 1342-1351.
- [S4] P. Fugedi, P. J. Garegg, *Carbohydr. Res.* **1986**, 149, C9-C12.
- [S5] W. Jahnke, H. C. Kolb, M. J. J. Blommers, J. L. Magnani, B. Ernst, *Angew. Chem. Int. Ed. Engl.* **1997**, 36, 2603-2607.
- [S6] a) S. Mesch, K. Lemme, H. Koliwer-Brandl, D. S. Strasser, O. Schwardt, S. Kelm, B. Ernst, *Carbohydr. Res.* **2010**, 345, 1348-1359; b) F. Bitsch, R. Aichholz, J. Kallen, S. Geisse, B. Fournier, J. M. Schlaeppli, *Anal. Biochem.* **2003**, 323, 139-149.
- [S7] a) G. Thoma, J. T. Patton, J. L. Magnani, B. Ernst, R. Oehrlein, R. O. Duthaler, *J. Am. Chem. Soc.* **1999**, 121, 5919-5929; b) G. Weitz-Schmidt, K. W. Gong, C. H. Wong, *Analytical Biochemistry* **1999**, 273, 81-88; c) G. Weitz-Schmidt, D. Stokmaier, G. Scheel, N. E. Nifant'ev, A. B. Tuzikov, N. V. Bovin, *Anal. Biochem.* **1996**, 238, 184-190.
- [S8] T. Wiseman, S. Williston, J. F. Brandts, L. N. Lin, *Anal. Biochem.* **1989**, 179, 131-137.
- [S9] a) J. Tellinghuisen, *Anal. Biochem.* **2008**, 373, 395-397; b) W. B. Turnbull, A. H. Daranas, *J. Am. Chem. Soc.* **2003**, 125, 14859-14866.
- [S10] W. S. Somers, J. Tang, G. D. Shaw, R. T. Camphausen, *Cell* **2000**, 103, 467-479.







Chapter 2.3. – Manuscript 1

E-Selectin Ligand Complexes Adopt an Extended High- Affinity Conformation

Manuscript in preparation.

Contributions

- | | |
|-----------------|---|
| R. C. Preston | <ul style="list-style-type: none">• Molecular cloning.• Generation of stably transfected CHO cell line.• Expression, deglycosylation, and protein purification.• Activity assay.• Small-angle X-ray scattering.• Protein crystallization, data collection, processing, and model building.• Manuscript preparation. |
| R. P. Jakob | <ul style="list-style-type: none">• Protein crystallization support. |
| F. P. C. Binder | <ul style="list-style-type: none">• Synthesis of sLe^x-OTMSE. |
| C. P. Sager | <ul style="list-style-type: none">• Computational calculations |
| T. Maier | <ul style="list-style-type: none">• Supervision of data collection, processing, and model building.• Manuscript preparation. |
-

E-Selectin Ligand Complexes Adopt an Extended High-Affinity Conformation

Roland C. Preston,[†] Roman P. Jakob,[‡] Florian P. C. Binder,[†] Christoph P. Sager,[†] Beat Ernst,^{*,†} and Timm Maier^{*,‡}

[†]Institute of Molecular Pharmacy, University of Basel, Klingelbergstrasse 50, 4056 Basel (Switzerland)

[‡]Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel (Switzerland)

ABSTRACT: E-selectin is a cell-adhesion molecule of the vascular endothelium that promotes essential leukocyte rolling in the early inflammatory response by binding to glycoproteins containing the tetrasaccharide sialyl Lewis^x (sLe^x). Efficient leukocyte recruitment under vascular flow conditions depends on an increased lifetime of E-selectin/ligand complexes under tensile force in a so-called catch-bond binding mode. Co-crystal structures of a representative fragment of the extracellular E-selectin region with sLe^x and a glycomimetic antagonist thereof reveal an extended E-selectin conformation, which is identified as a high-affinity binding state of E-selectin by molecular dynamics simulations. Small-angle X-ray scattering experiments demonstrate a direct link between ligand binding and E-selectin conformational transition under static conditions in solution. This permits tracing a series of concerted structural changes connecting ligand binding to conformational stretching as the structural basis of E-selectin catch-bond mediated leukocyte recruitment. The detailed molecular view of the binding site paves the way for the design of a new generation of selectin antagonists. This is of special interest, since their therapeutic potential was recently demonstrated with the pan-selectin antagonists GMI-1070 (Rivipansel).

INTRODUCTION

E-, P- and L-selectin are a family of Ca²⁺-dependent C-type lectins mediating cell-cell adhesion in the vascular system with a pivotal role in inflammation.¹ They are expressed on the vascular endothelium (E-, P-selectin), platelets (P-selectin) or leukocytes (L-selectin) and bind to cell-surface glycoproteins e.g. P-selectin glycoprotein ligand-1 (PSGL-1) or E-selectin ligand-1 (ESL-1) presenting the minimal tetrasaccharide binding motif sialyl Lewis^x (sLe^x, Figure 1a).²⁻⁴ Selectins exhibit catch-bond behavior: The lifetime of complexes of selectins with their ligands increases under tensile force-conditions⁵ as demonstrated first by atomic force microscopy and flow chamber assays for P-selectin,^{6,7} later also for L-selectin⁸⁻¹¹ and most recently for E-selectin.^{12,13} For E- and P-selectin the catch-bond binding mode is required under flow conditions for efficient leukocyte tethering and rolling along the vascular endothelium, a prerequisite for subsequent firm adhesion and transmigration to sites of inflammation.¹⁴ Due to their essential role in leukocyte recruitment, selectins are an attractive target for therapeutic intervention against diseases involving excessive inflammatory response as found in numerous cardiovascular and autoimmune diseases.¹⁵ A first demonstration of the therapeutic potential of selectin inhibition was very recently achieved by successful completion of phase-II clinical trials of the pan-selectin antagonists GMI-1070 (Rivipansel) for the treatment of the vaso-occlusive crisis in sickle-cell disease.^{16,17} The success of GMI-1070 now paves the way for the development of selectin-type specific antagonists for further indications.

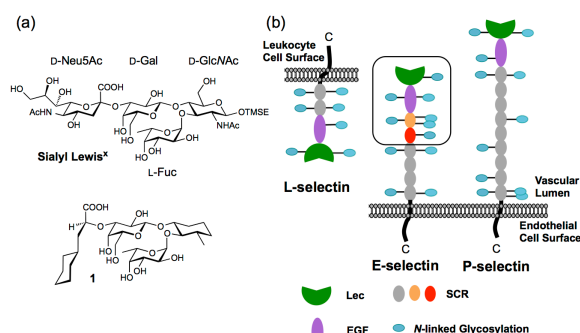


Figure 1. E-selectin ligands and selectin domain arrangement. (a) Chemical structures of sialyl Lewis^x (sLe^x) and glycomimetic **1**.¹⁸ (b) Comparison of human L-, E- and P-selectin. The representative E-selectin fragment used in this publication (E-selectin*) is boxed and domains are colored red, orange (SCR domains), magenta (EGF) and green (lectin) as in following figures.

The selectins share a conserved domain architecture (Figure 1b) with an N-terminal lectin domain (Lec), which contains the carbohydrate-binding site, an epidermal growth factor-like domain (EGF), variable numbers of short consensus repeat (SCR) domains, a single-span transmembrane region, and a C-terminal cytoplasmic domain. Isolated Lec-EGF domain fragments are sufficient for ligand binding,¹⁹ however, as demonstrated in cell-based assays, their affinity improves with an increasing number of attached SCR domains.²⁰ Several models have been established to link the catch-bond phenomenon to the structure of selectins and their ligands. In the *sliding-rebinding model*, the alignment of the binding interface to the tensile force allows the ligand to slide along the protein surface and continuously form new interactions.^{9,21}

Other models build upon an allosteric relationship between force-induced domain separation and conformational changes in the binding site.^{22–24} In the *one-state-two-pathway model*, a single low energy state exists under no-force conditions with a high rate of unbinding. When force is applied, the protein switches into a conformation with a decreased rate of unbinding, thus allowing for longer lifetimes of the interactions.²⁵ In the *two-state-two-pathway model*, high and low affinity conformations are in equilibrium under no-force conditions and the ligand can dissociate from either conformational state, albeit at different rates.²⁶ Under force conditions, the high-affinity state is preferred, leading to decreased unbinding up to a certain force threshold.

Currently, a single conformational state of the Lec-EGF didomain has been visualized for apo-P and E-selectin as well as for sLe^x-bound P- and E-selectin in crystallographic soaking experiments.^{27,28} However, the crystal structure of the P-selectin Lec-EGF didomain co-crystallized with a large phosphorylated glycopeptide fragment of its physiological glycoprotein ligand PSGL-1, yielded a second, more extended conformation suggesting a mechanism of conformational coupling specific for the binding of P-selectin to its multivalent peptide ligand.²⁷ It remains unclear how catch-bond behavior is mediated in E- and L-selectin and whether small molecule ligands are capable of mediating conformational transitions. To analyze the crosstalk between conformational changes and binding of the natural tetrasaccharide epitope sLe^x and the glycomimetic **1** (Figure 1a) to E-selectin, we used a representative four-domain fragment (E-selectin*) of the extracellular region of E-selectin comprising the Lec- and EGF-like domains as well as two SCR domains. While sLe^x has a weak affinity for E-selectin (K_D 878 μ M), glycomimetic **1** exhibits an enhanced affinity in the low micromolar range (K_D 19 μ M).²⁹ Here, the sialic acid and *N*-acetyl-D-glucosamine (GlcNAc) are replaced by (*S*)-cyclohexyl lactic acid and (1*R*,2*R*,3*S*)-3-methylcyclohexane-1,2-diol, respectively. Combining structural information from co-crystallizing E-selectin* with sLe^x or glycomimetic **1** with molecular dynamics simulations and solution small-angle X-ray scattering, we reveal a novel high-affinity binding state relevant for catch-bond interactions in E-selectin and provide the basis for future rational drug design of potent E-selectin antagonists.

METHODS

Ligand synthesis. The glycomimetic **1** {(1*R*,2*R*,3*S*)-2-[(α -L-fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl} 3-*O*-[sodium (1*S*)-1-carboxy-2-cyclohexyl-ethyl]- β -D-galactopyranoside was synthesized as reported previously.¹⁸ The tetrasaccharide trimethylsilyl ethyl sialyl Lewis^x (sLe^x-OTMSE) was synthesized as described in the Supporting Information (SI).

Protein expression and purification. The coding sequences of the human E-selectin Lec-EGF-SCR1+2 domains (E-selectin*) cDNA was cloned into the pcDNA3.1(+) expression vector (Life Technologies Ltd, UK) using standard procedures. Stably transfected CHO-K1 cells (ATCC-No. CCL-61) were generated using FuGENE[®] HD (Roche Applied Science, Switzerland) transfection, G418 treatment, and subsequent clone selection. Protein expression was carried out in monolayer cultures in the presence of kifunensine. Purification using the functional anti-E-selectin antibody 7A9 coupled to sepharose was carried out as described previously.²⁹ Concentrated high-mannosylated E-selectin was partially deglycosylated with endoglycosidase H yielding pro-

tein with seven single GlcNAc moieties attached to each *N*-linked glycosylation site. For details see SI.

Co-crystallization and structure determination. Deglycosylated E-selectin* was co-crystallized with sLe^x-OTMSE or glycomimetic **1** in a sitting drop setup in the presence of 0.2 M CaCl₂ using polyethylene glycol 8000 as precipitant. Plate-like crystals appeared within one day after seeding. The structures of E-selectin*/**1** and E-selectin*/sLe^x were determined by molecular replacement and refined using BUSTER.³⁰ For details on data collection and refinement see Table S1 and SI.

Small angle X-ray scattering (SAXS). SAXS measurements were conducted at beamline x12sa-cSAXS of the Swiss Light Source using partially deglycosylated E-selectin* in presence or absence of 3 mM glycomimetic **1**. Experimental data were processed with PRIMUS³¹ and theoretical scattering curves were evaluated with CRY SOL.³² A model for the E-selectin* bent conformation was generated using the coordinates of the apo-form of the E-selectin Lec-EGF didomain²⁸ combined with the coordinates of the two SCR domains from E-selectin*/**1**. For details see SI and Table S2.

Molecular dynamics simulations. Molecular dynamics simulations were carried out using the OPLS 2005 force field at a constant temperature of 300 K. Energetic and structural data were recorded in 4.8 ps intervals. From the collected frames, every tenth (1000 frames in total) was used to calculate the average ΔG applying the molecular mechanics generalized Born surface area (MM-GBSA) method. The details are described in SI and in Tables S3 and S4.

RESULTS

To analyze the implications of ligand binding to E-selectin, we determined the crystal structures of a representative fragment of the E-selectin extracellular region (E-selectin*) with either sLe^x or the glycomimetic **1**. Secreted E-selectin* expressed in CHO cells was purified by affinity chromatography with antibody 7A9, which blocks E-selectin functionality.³³ The purified protein after deglycosylation (Figure S1) contains single GlcNAc residues attached to each of the seven predicted *N*-glycosylation sites and is functional in a ligand binding assay (Figure S2). E-selectin* co-crystals with sLe^x or glycomimetic **1** were obtained in crystallization conditions based on polyethylene glycol 8000 in the presence of 0.2 M CaCl₂ and diffracted to 2.40 Å and 1.93 Å, respectively. Structure determination was carried out by molecular replacement, final structures were refined to R_{work}/R_{free} 0.217/0.253 and 0.183/0.232 for sLe^x and glycomimetic **1** co-crystals, respectively, with two virtually identical molecules per asymmetric unit (Table S1).

E-selectin ligand complexes adopt an extended linear conformation. Both ligand complexes of E-selectin* adopt an extended shape with a total length of 135 Å (Figure 2a). The binding modes of the physiological sLe^x ligand and glycomimetic **1** are almost identical (Figure 3a); both ligands bind into the same deep binding pocket covered by the loop 83-88, and their fucose moieties coordinate the central Ca²⁺ ion of the Lec-domain. However, compared to previously published structures of E-selectin Lec-EGF didomain fragments,^{27,28} either without ligand or with sLe^x soaked into preformed protein crystals, considerable rearrangements in the ligand binding site and a transition from a bent to an extended shape of the Lec-EGF domains are observed (Figure 3b). The SCR domains, which have not been previously visualized in the context

of selectin Lec-EGF domains, align in both ligand complexes in a linear fashion with the EGF domain (Figure 2b). Since E-selectin*/1 crystals diffract to higher resolution, the discussion focuses on this complex.

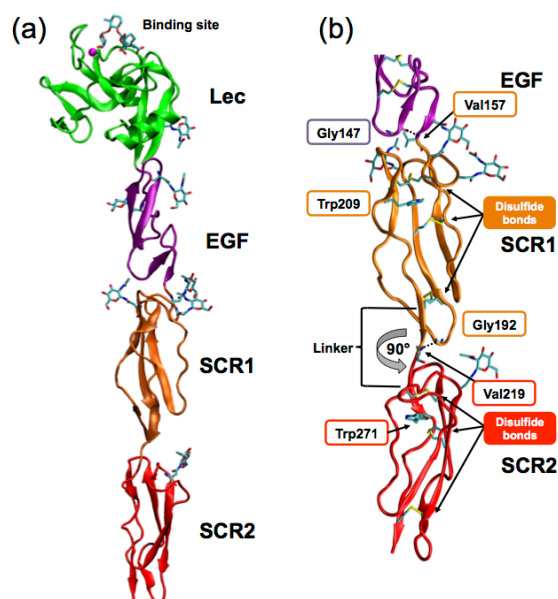


Figure 2. E-selectin short consensus repeat domains are linear rigidified domains. (a) Structure of E-selectin with the N-terminal Lec, the EGF-like- and the first two SCR domains. The glycomimetic **1** coordinates the Ca²⁺ ion (magenta) in the carbohydrate recognition site. Seven N-linked GlcNAc moieties are located at interdomain interfaces. (b) The SCR domains are aligned, with a 90° rotation of SCR2 relative to SCR1. Three disulfide bonds stabilize each SCR domain. Conserved tryptophans (Trp209/271) are wedged between the upper and central disulfide bond. SCR domain flexibility is limited by hydrogen bonds between Gly147 and Val157 or Gly192 and Val219.

Rigidified short consensus repeats at the basis of the selectin extracellular domain. In selectins, between two and nine SCR-domains provide the coupling of the Lec-EGF didomain to the cell surface. E-selectin comprises six SCR domains, two of which are included in E-selectin*. The sequence conservation of selectin SCR domains is low (Figure S3), however, they all consist of a 60 amino acid compact β -sheet fold (Figure 2b) and contain six conserved cysteines, two more than most other SCR domains (Table S5 and references therein). These cysteines form three intra-domain disulfide bonds, two common ones at each end of the domain, and an additional disulfide bond in the center. A conserved tryptophan (SCR1: Trp209, SCR2: Trp271) is interlocked between the N-terminal and the central disulfide bond (Figure 2b) resulting in the formation of a tightly packed rigid core in selectin SCR domains. The two successive SCR domains are rotated by 90° relative to each other. This orientation is stabilized by an interaction of the loop Gly192 backbone oxygen with the backbone NH of the linker region Val219. A similar stabilizing interaction is observed at the EGF/SCR1 interface between Gly147 and Val157. The linker between the linearly arranged E-selectin SCR1-2 domains is composed of only four residues (Asn217, Val218, Val219, and Glu220), while linker lengths of five to eight residues have been observed in non-linearly arranged SCR domains.³⁴ In contrast

to SCR domain linkers in other proteins, the pattern of hydrophilic-hydrophobic-hydrophobic-hydrophilic amino acids is well conserved in all selectin SCR linkers, except between SCR7 and 8 in P-selectin (Figure S3), indicating a critical role of linker composition and stabilizing interdomain interactions for selectin function.

The first E-selectin SCR domain possesses three N-linked glycosylation sites (Asn158, Asn178, Asn182), while the second SCR2 has a single site (Asn244). The observed glycosylation sites are located at the interface of the EGF-like/SCR1 and SCR1/SCR2 domains. While SCR4 and SCR5 have no predicted glycosylation sites, SCR3 and SCR6 each carry two at position 291, 311 (SCR3) and 482, 506 (SCR6). Based on the SCR1/2 structure, glycosylation at residue 291 is located at the interface to SCR2, and both sites in SCR6 are in close proximity to the membrane surface. This arrangement indicates that glycosylation generally contributes to the stabilization of SCR domain linking regions. Together with the conserved tryptophans, the additional disulfide bond and the conserved linker region, glycosylation thus contributes to the formation of a reinforced SCR stalk for mechanosensing by the Lec-EGF didomain.

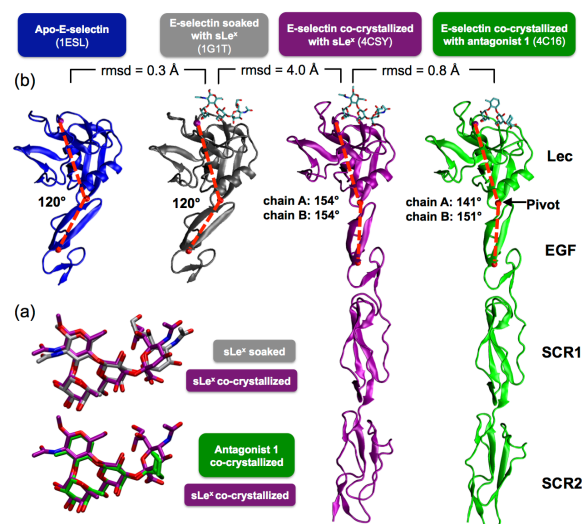


Figure 3. E-selectin adopts an extended overall conformation upon ligand binding. (a) Superposition of sLe^x soaked into preformed crystals²⁷ (gray, upper panel) and the co-crystallized glycomimetic **1** (green, lower panel) onto the co-crystallized sLe^x (purple). (b) Apo-E-selectin (blue, PDB code 1ESL)²⁸ and E-selectin soaked with sLe^x (gray, PDB code 1G1T)²⁷ adopt a bent conformation. The extended state observed in co-crystals with sLe^x (magenta) or glycomimetic **1** (green) is characterized by a separation of the Lec- and EGF-like domains. Root-mean-square deviations (rmsd) were calculated for Lec- and EGF-like domains only. The ligand binding Ca²⁺ ion is depicted as a magenta sphere, C α atoms of Trp1 (pivot) and Cys144 (EGF-like domain) as red spheres.

Ligand binding induces the extended selectin conformation. The relative orientation of the Lec- and EGF-domains observed here in both ligand complexes is considerably different from the bent conformation observed previously in crystals of a E-selectin Lec-EGF didomain fragment without ligand or soaked with sLe^x.^{27,28} In the bent conformation, the angle between the Lec- and EGF domains is ~120°, while it is 141° and 151° in the current co-crystal structures (Figure 3b, Video S1 available as web enhanced

object). This bent-to-extended movement occurs by a rotation around a pivot at the Lec-EGF interface and results in a separation of the two domains. Small-angle X-ray scattering (SAXS) of E-selectin* demonstrates the impact of ligand binding on E-selectin in solution. In the absence of ligand, scattering curves are in good agreement with a bent model obtained by combining the apo-E-selectin structure²⁸ with the SCR domains from the current E-selectin* structures (Figure 4a). Scattering curves in the presence of glycomimetic **1** reveal a significant shift at $q = 0.3\text{--}0.4 \text{ \AA}^{-1}$ (Figure 4b), which correlates well with calculated curves for the extended conformation. These results demonstrate that the solution conformation closely resembles the conformation in the crystal structures. Ligand binding therefore indeed shifts the conformational equilibrium of E-selectin from the bent to the extended state, while the linear arrangement of SCR domains is maintained irrespective of the presence or absence of the ligand.

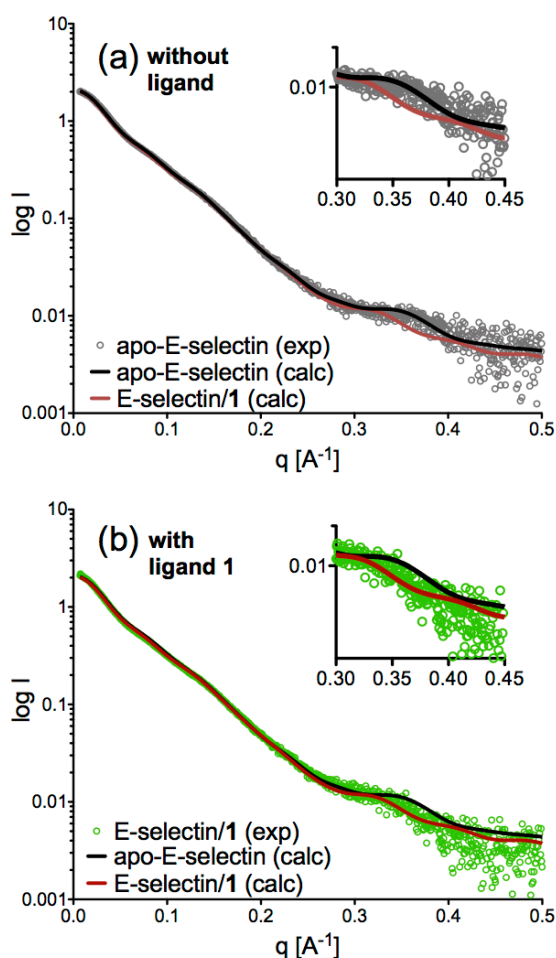


Figure 4. In solution small-angle X-ray scattering analysis reveals the conformational differences between apo-E-selectin and the E-selectin/**1** complex. (a, b) Calculated (black or red lines) and experimental (circles) scattering curves of E-selectin in absence (gray circles) or presence (green circles) of the ligand. (a) In absence of glycomimetic **1**, E-selectin adopts the bent conformation. (b) In presence of the ligand, the experimental scattering curve shows a significant shift at $q = 0.3\text{--}0.4 \text{ \AA}^{-1}$ corresponding to a transition towards the extended conformation.

The extended conformation is a high-affinity state.

Calculated binding free energies in a combined approach using molecular dynamics simulations with an explicit solvation environment and a subsequent scoring of an ensemble of MD frames with the MM-GB/SA method demonstrate that additional interactions of protein and ligand in the extended conformation result in higher interaction energies in comparison to the shallow binding site in the bent state (Table S4). This supports the notion that the extended conformation represents a high-affinity binding state of E-selectin.

Alterations in the lectin domain connect ligand binding to conformational change.

Three regions located in the Lec domain are affected by conformational changes upon ligand binding: The binding site loop 81-89 changes its conformation (Figure 5a), and a bridging region formed by two loops (residues 52-75) transduces this motion (Figure 5b) to the pivot region consisting of the pivot residue Trp1 (Figure 5c) and surrounding residues in the Lec (residues 22-23 and 30-31, Figures 5c and 5d) and in the EGF-like domain (Asn138). These variable regions are mounted onto a rigid scaffold of the lectin domain, stabilized by two disulfide bonds (Cys19-Cys117 and Cys90-Cys109). The alterations extend over 35 Å from the carbohydrate-binding site to the interface with the EGF-like domain.

The ligand binding site. The most extensive conformational changes are observed within the binding site loop 81-89 (Figure 5a), where the Ca atom of residue Gln85 moves by 10 Å (see Video S2 available as web enhanced object). This leads to a constriction around the central Ca^{2+} ion and the L-fucose (L-Fuc) moiety of the ligand, which coordinates the Ca^{2+} ion with its 3- and 4-hydroxyl groups. In the bent conformation Asn83 coordinates the Ca^{2+} ion, while in the extended conformation Glu88 swings inwards to coordinate the Ca^{2+} ion and displaces Asn83 towards the solvent. This swapping is supported by an additional interaction between Glu88 and the 2-hydroxyl group of L-Fuc. Gln85, which is solvent-exposed in the bent conformation, also shifts towards the ligand into hydrogen bonding distance to the 2-hydroxyl group. Thus, as a consequence of the conformational adaptation, the 2-hydroxyl group of L-Fuc directly interacts with the protein in the extended conformation, while it is only involved in water-mediated interactions to Glu107 and Asn83 in the bent conformation. These findings explain the importance of the L-Fuc 2-hydroxyl group for ligand affinity,^{35,36} which was not rationalized by earlier structural data.²⁷ Finally, also the loop around Pro46 shifts from an original distance of 6.8 Å to the 6-hydroxyl group of the D-galactose (D-Gal) moiety in the bent conformation towards 4.5 Å in the extended conformation, where the backbone oxygen of Pro46 mediates a water-bridged interaction to the 6-hydroxyl group of D-Gal.

The bridging region. The constriction around the L-Fuc binding site leads to the disruption of interactions at the interface of the binding site and the bridging loops 52-75 (Figure 5b) between the Asp89 and Arg54 side chains and Lys55 backbone, respectively. The bridging region itself shifts in a rigid body motion from the bent to the extended conformation. This rigidity is maintained by backbone interactions between Arg54 and Trp60/Val61, Gly52 and Val63 as well as Trp60 and Leu69. The rigid body motion serves as an allosteric bridge for transmitting the conformational alteration from the binding site to the pivot region.

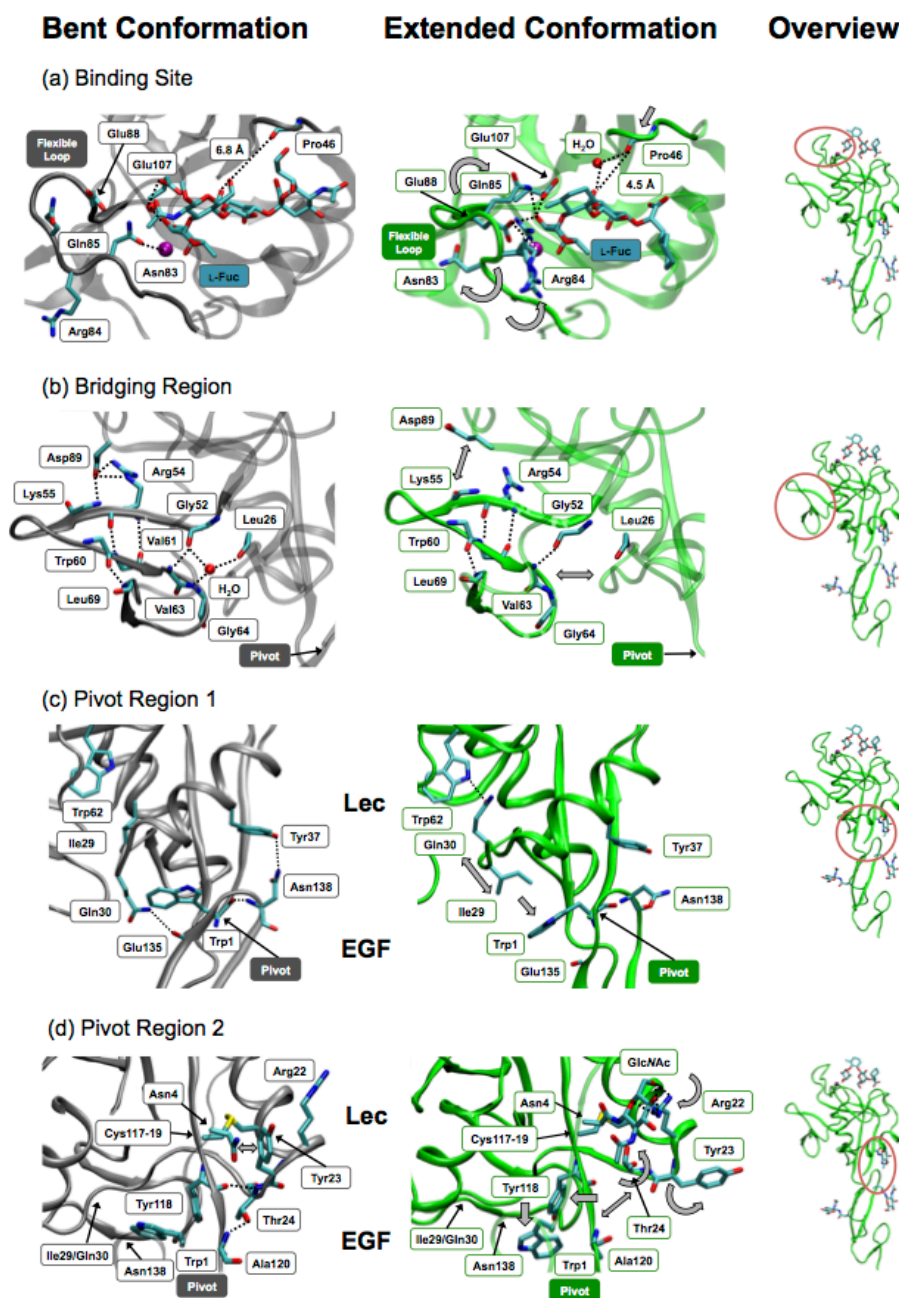


Figure 5. The bent and extended conformation of E-selectin exhibit differences in the ligand binding site reaching to the Lec-EGF domain interface. (a) The binding site is constricted by rearrangements of loop 81–89. Glu88 replaces Asn83 as Ca^{2+} coordinating residue and forms an additional interaction with the 2-hydroxyl group of L-Fuc. (b) The bridging region 52–70 undergoes a rigid body motion breaking the interactions of Arg54/Lys55 with Asp89 of the binding site. At the pivot interface, a water mediated hydrogen bond network involving Gly52, Val63 and Leu26 is perturbed. (c) Within the pivot region, Ile29 and Gln30 flip orientation, which disrupts interdomain interactions to Glu135 and favors a dislocation of Trp1, leading to the hinge-bending motion between the Lec and EGF domains. The Tyr37–Asn138 interdomain interaction is broken. (d) Another shift in the pivot region originates at the GlcNAc residue linked to Asn4. The interaction of Arg22 with the carbohydrate further disrupts interdomain interactions and releases the Tyr118 side chain towards the pivot residue Trp1.

The pivot region. The E-selectin bent conformation is stabilized by an extensive hydrogen bond network between the Lec- and EGF-like domains with main contributions from Trp1, Thr24, Gln30, Tyr37, Ala120, Glu135, and Asn138 (Figure 5c

and 5d). The Trp1 hinge residue is the only residue of the Lec domain that coherently moves with the EGF-like domain to the extended conformation. This motion is accompanied by (i) a flip of residues Ile29/Gln30 (Figure 5c), (ii) a flip of residues

Tyr23/Thr24 accompanied by an altered side chain conformation of Arg22 (Figure 5d) and (iii) a disruption of the interdomain interaction between Tyr37/Asn138 (Figure 5c). A flip of the residues Ile29 and Gln30, located in a solvent exposed loop with increased flexibility between helix 1 and 2 of the lectin domain, facilitates the movement of Trp1 (Figure 5c). In the bent conformation, Ile29 is enclosed in a hydrophobic pocket, while Gln30 is involved in intra- and interdomain interactions to Glu34 and Glu135. When the extended conformation is adopted, the steric hindrance of Trp1 on Ile29 and Gln30 is released and the flipped conformation of Gln30 is stabilized by new interactions with the backbone oxygen of Gly102 and the side chain of Trp104. These rearrangements in the pivot region disrupt a contact to the bridging region mediated by a water molecule bridging the backbone of Leu26 to Gly52/Gly64 (Figure 5b). As a consequence, the pivot and bridging regions are further separated and additional water molecules enter into the crevice between the domains. Tyr23 and Thr24 are restrained in the bent conformation by the side chain of Arg22. While Arg22 stacks with Tyr18 in the bent conformation, its side chain is re-oriented in the extended conformation and interacts with the GlcNAc attached to Asn4. The movement of Arg22 provides space for a flip of residues Tyr23 and Thr24 of the pivot region (Figure 5d). Tyr23 relaxes from an unfavorable into a favorable position in the Ramachandran plot. The new conformation is stabilized by an interaction of the Thr24 side chain with the Cys19 backbone oxygen. The flip of Tyr23/Thr24 further disrupts interactions between the Lec- and the EGF-like domain, namely between Thr24 backbone and Tyr118 backbone and Thr24 side chain with Ala120. The side chain of Tyr118 enters into a cleft that is opened by the movement of Trp1. A hydrogen bond between Tyr37 and Asn138 contributes to the interface between the Lec and EGF-like domain in the bent conformation of E- and L-selectin (unpublished, PDB code 3CFW). In the E-selectin*/1 structure, the Asn138 residue is pushed towards the Lec domain, thus disrupting the interdomain hydrogen bond to Tyr37 (Figure 5c). This conformational change is additionally accompanied by a backbone flip between Ile137 and Asn138.

A transition from bent to extended conformation in selectins has previously been observed only for the Lec-EGF didomain of P-selectin co-crystallized with a multivalent PSGL-1 peptide ligand.²⁷ This peptide comprised the 19 N-terminal amino acids of mature PSGL-1, including an sLe^x-modified glycan attached to Thr16 and three sulfated tyrosine residues, which are directly involved in protein ligand interactions.^{27,37,38} Based on the P-selectin/PSGL-1 crystal structure, a P-selectin specific mechanism for conformational coupling was proposed.²³ Despite a sequence identity in the Lec- and EGF-like domains of 62 % between E- and P-selectin (Figure S4), two key differences between E- and P-selectin supported the P-selectin-specificity of the proposed mechanism: First, E-selectin does not interact with sulfated tyrosines, while those mediate key protein-ligand interaction in the P-selectin/PSGL-1 complex.^{39,40} Second, differences in the Lec-EGF domain interface suggested a stabilization of the bent conformation in E-selectin by an additional hydrogen bond.^{27,28} The current data demonstrate that E- and P-selectin undergo a similar overall conformational change, but with considerable differences in the mechanism of conformational coupling. (i) The E-selectin Lec-EGF-domain interface is locked by a hydrogen bond between Asn138 and Tyr37. This interaction is

absent in P-selectin, which carries a glycine at position 138. The functional relevance of this interdomain interaction was demonstrated on L-selectin, where an Asn138Gly mutant improved ligand binding under hydrodynamic force.^{9,41} It has been hypothesized, that Asn138 would be forced into an unfavorable backbone conformation upon disruption of the Tyr37-Asn138 hydrogen bond, which would finally prevent the transition into the extended state.²³ The backbone flip between residues 137 and 138 observed here, now demonstrates how such non-favored conformation is avoided. (ii) While the Trp1 residue has also been identified as the pivot for P-selectin, the flip of the Ile29 and Gln30 side chains was not observed. (iii) Residue Tyr23 in the direct vicinity of the pivot residue Trp1 is unaltered in the bent and extended conformation of P-selectin, while it relaxes from an unfavorable into a favorable backbone conformation for E-selectin.

DISCUSSION

The catch-bond behavior of selectins^{6,7,12,13} is crucial for their ability to mediate leukocyte rolling and is achieved by a transition between a bent and an extended conformational state. This has been elegantly demonstrated by forcing P-selectin into an extended conformation by modifications at the Lec-EGF interface: A P-selectin variant locked in an extended conformation displays enhanced ligand binding under static conditions, but its ability to mediate leukocyte rolling under flow conditions was impaired.^{7,41} However, an extended conformation has so far only been observed for P-selectin in complex with a glycopeptide ligand that interacts with its carbohydrate epitope as well as three sulfated tyrosines. Here, we reveal an extended state of E-selectin in complex with its minimal binding motif sLe^x. Based on an analysis of protein-ligand interactions and quantitative analysis of molecular dynamics simulations, this extended conformation represents a true high affinity ligand binding state of E-selectin. SAXS measurements confirm that ligand binding induces a conformational transition from the bent low affinity to the extended high-affinity state in solution and thus demonstrate that glycoprotein ligands such as PSGL-1 or ESL-1 are not strictly required for efficiently promoting conformational changes in selectins. Even a small molecular ligand such as sLe^x in the absence of flow conditions is sufficient to initiate a structural transition to an extended conformation.

To analyze the functional mechanisms underlying catch-bond behavior, detailed structural knowledge of all relevant states of E-selectin is required (Figure 6). Apo-E-selectin adopts a bent conformation of the Lec-EGF-domains as demonstrated by crystallographic analysis²⁸ and SAXS data. The bent conformation is capable of recognizing and weakly binding sLe^x, as demonstrated in a crystallographic ligand soaking experiment, where crystal packing restraints prevented the transition to the extended conformation.²⁷ However, in solution ligand binding results in a conformational transition into the extended high-affinity binding state revealed here by co-crystallization and SAXS. The conformational change from a bent to an extended conformation allows E-selectin to support fast binding of rapidly flowing leukocytes to the vascular endothelium at the leading edge of the cell while controlling the rate of unbinding at the trailing edge, thus providing mechanical stability to the rolling of leukocytes along the vasculature. Under *in vivo* flow conditions, the extended conformation is favored by tensile force pulling

along the E-selectin-ligand axis reinforcing the extended conformation and thus increasing lifetime of the interaction. The rigidity of the linearly arranged SCR domains is promoted by conserved interdomain contacts and three intradomain disulfide bonds and provides a basis for mechanosensing by the Lec-EGF didomain of E-selectin. Our findings provide direct evidence for the *two-state-two-pathway model* of conformational changes underlying catch-bond properties in E-selectin (Figure 6). The bent conformation is prevailing in solution and recognizes ligands in a low-affinity encounter complex. The encounter complex undergoes a conformational change to a high-affinity extended state already under no-force conditions, but tensile force acting along the Lec-EGF-ligand binding site axis further increases its lifetime.

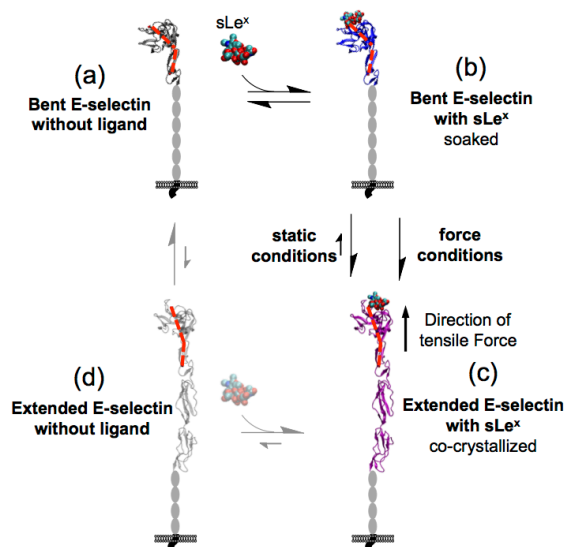


Figure 6. The ligand-dependent equilibrium between bent and extended conformation of E-selectin. In absence of ligand, E-selectin primarily adopts the bent conformation (a) and not an extended conformation (d). The bent conformation is capable of ligand binding (b), although with low affinity. Upon ligand binding E-selectin favors the extended high-affinity state (c). Tensile force along the E-selectin-ligand axis exerted by tethered leukocytes under flow conditions increase the lifetime of the high affinity conformation.

The ligand binding site undergoes dramatic conformational changes between the low-affinity and the high-affinity state (Video S1 available as web enhanced object). The shallow binding surface is transformed into a pronounced binding pocket with multiple additional protein-ligand interactions, while the conformation of the ligand remains unchanged. Co-crystallization with the glycomimetic **1** demonstrates the relevance of the high-affinity conformation for the design of drug candidates targeting E-selectin. The detailed knowledge into ligand recognition of E- and P-selectin, opens the way towards the rational design of novel E-selectin specific antagonists. Based on the success of the pan-selectin antagonist Rivipansel, E-selectin specific antagonists are promising drug candidates that may further target specific indications related to excessive inflammatory responses in human disease. In summary, the current results contribute to a comprehensive understanding of the mechanism of mechanosensing and catch-bond behavior in the selectin family and they reveal the

pathway of ligand-induced conformational changes in E-selectin with critical relevance for ongoing targeted drug development.

ASSOCIATED CONTENT

Supporting Information

Expanded Materials and Methods including sLe^x-OTMSE synthesis, protein purification, expression, and activity assays, protein-ligand co-crystallization, structure determination and refinement, SAXS analysis, and molecular dynamics simulations. Supporting tables and figures including Lec-, EGF-, and SCR domain analysis.

Web Enhanced

Videos S1 and S2 are available as web enhanced objects in the HTML version of the paper.

AUTHOR INFORMATION

Corresponding Authors

tim.m.maier@unibas.ch; beat.ernst@unibas.ch

Notes

The authors declare no competing financial interests.

ACKNOWLEDGMENT

The authors thank the staff of beamline PXIII and Andreas Menzel from the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland.

REFERENCES

- (1) Lasky, L. A. *Annu. Rev. Biochem.* **1995**, *64*, 113.
- (2) Walz, G.; Aruffo, A.; Kolanus, W.; Bevilacqua, M.; Seed, B. *Science* **1990**, *250*, 1132.
- (3) Phillips, M. L.; Nudelman, E.; Gaeta, F. C.; Perez, M.; Singhal, A. K.; Hakomori, S.; Paulson, J. C. *Science* **1990**, *250*, 1130.
- (4) Polley, M. J.; Phillips, M. L.; Wayner, E.; Nudelman, E.; Singhal, A. K.; Hakomori, S.; Paulson, J. C. *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 6224.
- (5) Thomas, W. E.; Vogel, V.; Sokurenko, E. *Annual review of biophysics* **2008**, *37*, 399.
- (6) Marshall, B. T.; Long, M.; Piper, J. W.; Yago, T.; McEver, R. P.; Zhu, C. *Nature* **2003**, *423*, 190.
- (7) Waldron, T. T.; Springer, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 85.
- (8) Zhu, C.; Yago, T.; Lou, J.; Zarnitsyna, V. I.; McEver, R. P. *Ann. Biomed. Eng.* **2008**, *36*, 604.
- (9) Lou, J. Z.; Yago, T.; Klopocki, A. G.; Mehta, P.; Chen, W.; Zarnitsyna, V. I.; Bovin, N. V.; Zhu, C.; McEver, R. P. *J. Cell Biol.* **2006**, *174*, 1107.
- (10) Finger, E. B.; Puri, K. D.; Alon, R.; Lawrence, M. B.; von Andrian, U. H.; Springer, T. A. *Nature* **1996**, *379*, 266.
- (11) Ramachandran, V.; Nollert, M. U.; Qiu, H. Y.; Liu, W. J.; Cummings, R. D.; Zhu, C.; McEver, R. P. *Proc. Natl. Acad. Sci. U. S. A.* **1999**, *96*, 13771.
- (12) Wayman, A. M.; Chen, W.; McEver, R. P.; Zhu, C. *Biophys. J.* **2010**, *99*, 1166.
- (13) Snook, J. H.; Guilford, W. H. *Cellular and molecular bioengineering* **2010**, *3*, 128.
- (14) Kansas, G. S. *Blood* **1996**, *88*, 3259.
- (15) Ley, K. *Trends Mol. Med.* **2003**, *9*, 263.
- (16) Pfizer Inc., GlycoMimetics Inc.; Study of GMI-1070 for the Treatment of Sickle Cell Pain Crisis.: Clinicaltrials.gov [Internet]. Bethesda (MD) National Library of Medicine (US). 2000- [cited 2014 Sep 22]. Available from: <https://clinicaltrials.gov/ct2/show/NCT01119833> NLM Identifier NCT01119833.
- (17) Pfizer Inc., GlycoMimetics Inc.; Study of Intravenous GMI-1070 in Adults With Sickle Cell Disease.: Clinicaltrials.gov [Internet]. Bethesda (MD) National Library of Medicine (US). 2000- [cited 2014 Sep 22]. Available from: <https://clinicaltrials.gov/ct2/show/NCT00911495> NLM Identifier NCT00911495.

-
- (18) Schwizer, D.; Patton, J. T.; Cutting, B.; Smiesko, M.; Wagner, B.; Kato, A.; Weckerle, C.; Binder, F. P.; Rabbani, S.; Schwardt, O.; Magnani, J. L.; Ernst, B. *Chemistry* **2012**, *18*, 1342.
- (19) Erbe, D. V.; Wolitzky, B. A.; Presta, L. G.; Norton, C. R.; Ramos, R. J.; Burns, D. K.; Rumberger, J. M.; Rao, B. N. N.; Foxall, C.; Brandley, B. K.; Lasky, L. A. *J. Cell Biol.* **1992**, *119*, 215.
- (20) Li, S. H.; Burns, D. K.; Rumberger, J. M.; Presky, D. H.; Wilkinson, V. L.; Anostario, M., Jr.; Wolitzky, B. A.; Norton, C. R.; Familletti, P. C.; Kim, K. J.; et al. *J. Biol. Chem.* **1994**, *269*, 4431.
- (21) Lou, J. Z.; Zhu, C. *Biophys. J.* **2007**, *92*, 1471.
- (22) Thomas, W.; Forero, M.; Yakovenko, O.; Nilsson, L.; Vicini, P.; Sokurenko, E.; Vogel, V. *Biophys. J.* **2006**, *90*, 753.
- (23) Springer, T. A. *Proc. Natl. Acad. Sci. U. S. A.* **2009**, *106*, 91.
- (24) Yakovenko, O.; Sharma, S.; Forero, M.; Tchesnokova, V.; Aprikian, P.; Kidd, B.; Mach, A.; Vogel, V.; Sokurenko, E.; Thomas, W. E. *J. Biol. Chem.* **2008**, *283*, 11596.
- (25) Evans, E.; Leung, A.; Heinrich, V.; Zhu, C. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 11281.
- (26) Beste, M. T.; Hammer, D. A. *Proc. Natl. Acad. Sci. U. S. A.* **2008**, *105*, 20716.
- (27) Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000**, *103*, 467.
- (28) Graves, B. J.; Crowther, R. L.; Chandran, C.; Rumberger, J. M.; Li, S.; Huang, K. S.; Presky, D. H.; Familletti, P. C.; Wolitzky, B. A.; Burns, D. K. *Nature* **1994**, *367*, 532.
- (29) Binder, F. P.; Lemme, K.; Preston, R. C.; Ernst, B. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 7327.
- (30) Blanc, E.; Roversi, P.; Vonnhein, C.; Flensburg, C.; Lea, S. M.; Bricogne, G. *Acta Crystallogr D* **2004**, *60*, 2210.
- (31) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. *J Appl Crystallogr* **2003**, *36*, 1277.
- (32) Svergun, D.; Barberato, C.; Koch, M. H. J. *J Appl Crystallogr* **1995**, *28*, 768.
- (33) Rodriguez-Romero, A.; Almog, O.; Tordova, M.; Randhawa, Z.; Gilliland, G. L. *J. Biol. Chem.* **1998**, *273*, 11770.
- (34) Prota, A. E.; Sage, D. R.; Stehle, T.; Fingerioth, J. D. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 10641.
- (35) Ramphal, J. Y.; Zheng, Z. L.; Perez, C.; Walker, L. E.; Defrees, S. A.; Gaeta, F. C. A. *J. Med. Chem.* **1994**, *37*, 3459.
- (36) Brandley, B. K.; Kiso, M.; Abbas, S.; Nikrad, P.; Srivasatava, O.; Foxall, C.; Oda, Y.; Hasegawa, A. *Glycobiology* **1993**, *3*, 633.
- (37) Wilkins, P. P.; Moore, K. L.; McEver, R. P.; Cummings, R. D. *J. Biol. Chem.* **1995**, *270*, 22677.
- (38) Sako, D.; Comess, K. M.; Barone, K. M.; Camphausen, R. T.; Cumming, D. A.; Shaw, G. D. *Cell* **1995**, *83*, 323.
- (39) Rodgers, S. D.; Camphausen, R. T.; Hammer, D. A. *Biophys. J.* **2000**, *79*, 694.
- (40) Li, F. G.; Wilkins, P. P.; Crawley, S.; Weinstein, J.; Cummings, R. D.; McEver, R. P. *J. Biol. Chem.* **1996**, *271*, 3255.
- (41) Phan, U. T.; Waldron, T. T.; Springer, T. A. *Nat. Immunol.* **2006**, *7*, 883.
-

Supplementary Information

E-Selectin Ligand Complexes Adopt an Extended High-Affinity Conformation

Roland C. Preston,[†] Roman P. Jakob,[‡] Florian P. C. Binder,[†] Christoph P. Sager,[†] Beat Ernst,^{*,†} and Timm Maier^{*,‡}

[†]Institute of Molecular Pharmacy, University of Basel, Klingelbergstrasse 50, 4056 Basel (Switzerland)

[‡]Biozentrum, University of Basel, Klingelbergstrasse 50/70, 4056 Basel (Switzerland)

Contents

Synthesis of sialyl Lewis ^x -OTMSE	S2
Molecular cloning	S7
Protein purification	S7
Activity assays	S8
Protein-ligand co-crystallization	S9
Structure determination and refinement	S10
Small angle X-ray scattering	S11
Molecular dynamics simulations	S12
SCR comparison	S14
Lec-EGF comparison	S15
Supplementary References	S16

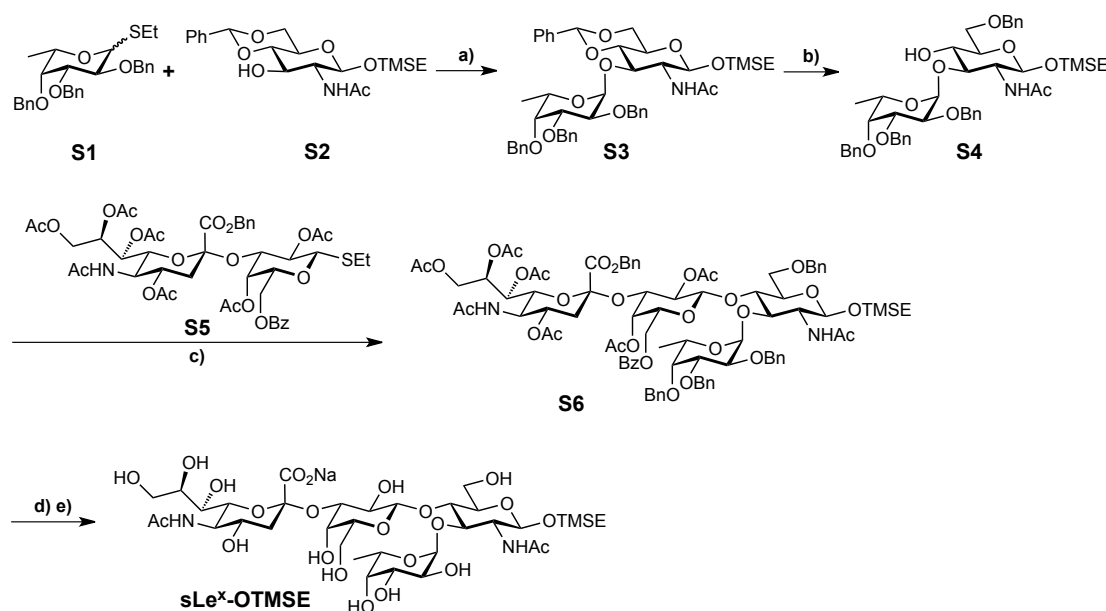
Synthesis of sialyl Lewis^x-OTMSE

General methods: NMR spectra were recorded on a Bruker Avance DMX-500 (500 MHz) spectrometer. Assignment of ¹H and ¹³C NMR spectra was achieved using 2D methods (COSY, HSQC, HMQC, HMBC). Chemical shifts are given in ppm and were assigned in relation to the solvent signals on the δ -scale¹ or to tetramethylsilane (0 ppm) as internal standard. Coupling constants *J* are given in Hertz (Hz). Multiplicities were specified as follows: s (singlet), d (doublet), dd (doublet of a doublet), t (triplet), q (quartet), m (multiplet). For assignment of resonance signals to the appropriate nuclei the following abbreviations were used: Fuc (fucose), Gal (galactose), GlcNAc (*N*-acetylglucosamine), Sia (sialic acid). Reactions were monitored by TLC using glass plates coated with silica gel 60 F₂₅₄ (Merck) and visualized by using UV light and/or by charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10 % H₂SO₄). Column chromatography was performed manually using silica gel 60 (40–63 μ m) from Fluka or using automated systems (RediSep Companion or RediSep Rf) from Teledyne Isco with normal phase RediSep columns from the same manufacturer or reversed-phase columns containing LiChroprep RP-18 (40–63 μ m) from Merck KGaA (Darmstadt, Germany). LC-MS separations were carried out using Sunfire C₁₈ columns (19 x 150 mm, 5.0 μ m) on a Waters 2525 LC, equipped with Waters 2996 photodiode array and Waters micromass ZQ MS for detection. Size exclusion chromatography was performed with Bio-Gel[®] P-2 Gel (45-90 mm) from Bio-Rad.

Solvents were purchased from Sigma-Aldrich or Acros. Solvents were dried prior to use where indicated. Tetrahydrofuran (THF) was dried by refluxing with sodium/benzophenone and distilled immediately before use. Dichloromethane (CH₂Cl₂) and dimethoxyethane (DME) were dried by filtration over Al₂O₃ (Fluka, type 5016 A basic). Methanol (MeOH) was dried by distillation from sodium methoxide, DMF by distillation from calcium hydride. Optical rotations were measured using a Perkin-Elmer Polarimeter 341. Electron spray ionization mass spectra (ESI-MS) were obtained on a Waters micromass ZQ. HRMS analysis were carried out using a Agilent 1100 LC equipped with a photodiode array detector and a Micromass QTOF I equipped with a 4 GHz digital-time converter. Microanalysis was performed at the Institute of Organic Chemistry at the University of Basel, Switzerland. Purity of final compounds was determined on an Agilent 1100 HPLC; detector ELS, Waters 2420; column: Waters Atlantis dC18, 3 mm, 4.6 x 75 mm; eluents: A: water + 0.1% TFA; B:

90% acetonitrile + 10% water + 0.1% TFA; linear gradient: 0-1 min 5% B; 1-20 min 5 to 70% B; flow: 0.5 mL/min.

Thiofucoside **S1**² and D-GlcNAc acceptor **S2**³ were coupled under *in situ* anomerization conditions to give disaccharide **S3**⁴, which upon regioselective opening of the benzylidene acetal⁵ yielded acceptor **S4** (Scheme S1). SLe^x precursor **S6** was synthesized *via* DMTST⁶ promoted coupling of **S5**⁷ and **S4**. Hydrogenolytic debenzoylation and saponification finally afforded TMSE protected sLe^x.



Scheme S1. a) Br₂, Bu₄NBr, CH₂Cl₂, MS 4 Å, 0 °C to r.t., 12 h, 55%; b) Me₃N·BH₃, AlCl₃, H₂O, THF, r.t., 5 h, 78%; c) DMTST, CH₂Cl₂, MS 4 Å, r.t., 5 d, 20%; d) H₂, Pd(OH)₂/C, dioxane, H₂O, r.t., 24 h; e) aq. NaOH, MeOH, r.t., 20 h, 43% from **S5**.

2-(Trimethylsilyl)ethyl 2,3,4-tri-*O*-benzyl- α -L-fucopyranosyl-(1 \rightarrow 3)-2-acetamido-4,6-*O*-benzylidene-2-deoxy- β -D-glucopyranoside (S3**).** Bromine (0.18 mL, 3.50 mmol) was slowly added to a stirred solution of donor **S1** (1.52 g, 3.18 mmol) in anhydrous CH₂Cl₂ (90 mL) at 0 °C under argon. After 10 min, excess bromine was quenched by addition of cyclohexene. Powdered activated molecular sieves 4 Å (9 g), TBAB (1.54 g, 4.78 mmol) and **S2**³ (0.65 g, 1.59 mmol) were added and the mixture was stirred at 0 °C for 2 h then warmed to r.t. and stirred overnight. The reaction mixture was filtered through celite and washed with water (100 mL) and brine (100 mL). The organic layer was dried over Na₂SO₄, filtered, and concentrated under reduced pressure. Column chromatography on silica (petroleum

ether/Et₂O) afforded **S3** (720 mg, 0.87 mmol, 55%) as white foam.

¹H NMR (500.1 MHz, CDCl₃): δ = 7.51-7.23 (m, 20H, Ar-H), 5.59 (d, J = 6.7 Hz, 1H, NH), 5.49 (s, CHPh), 5.04 (d, J = 3.2 Hz, Fuc-H1), 4.99 (d, J = 8.3 Hz, GlcNAc-H1), 4.96-4.88 (m, 2H, CH₂Ph), 4.75 (s, 2H, CH₂Ph), 4.65 (B of AB, J = 11.7 Hz, CH₂Ph), 4.58 (B' of A'B', J = 11.5 Hz, CH₂Ph), 4.39-4.31 (m, 2H, GlcNAc-H3, GlcNAc-H6a), 4.12-4.04 (m, 2H, Fuc-H2, Fuc-H5), 3.98-3.86 (m, 2H, Fuc-H3, CH₂CH₂Si), 3.76 (t, J = 10.0 Hz, GlcNAc-H6b), 3.63-3.45 (m, 4H, Fuc-H4, GlcNAc-H4, GlcNAc-H5, CH₂CH₂Si), 3.28 (m, 1H, GlcNAc-H2), 1.62 (s, 3H, COCH₃), 1.01-0.82 (m, 2H, CH₂CH₂Si), 0.80 (d, J = 6.4 Hz, Fuc-H6), 0.01 (s, 9H, Si(CH₃)₃); ¹³C NMR (125.8 MHz, CDCl₃): δ = 170.6 (COCH₃), 138.9, 138.8, 138.7, 129.2, 128.8, 128.5, 128.4, 128.3, 128.0, 127.9, 127.7, 127.5, 126.4 (24C, Ar-C), 101.8 (CHPh), 100.1 (GlcNAc-C1), 98.5 (Fuc-C1), 81.0 (GlcNAc-C4), 80.0 (Fuc-C3), 77.7, 77.4 (Fuc-C2, Fuc-C4), 75.0 (CH₂Ph), 74.9 (GlcNAc-C3), 74.3, 72.6 (CH₂Ph), 69.1 (GlcNAc-C6), 67.5 (CH₂CH₂Si), 66.9 (Fuc-C5), 66.3 (GlcNAc-C5), 59.1 (GlcNAc-C2), 23.4 (COCH₃), 18.2 (CH₂CH₂Si), 16.4 (Fuc-C6), -1.3 (3C, Si(CH₃)₃); [α]_D -73.3 (c 0.96, CHCl₃); MS (ESI): m/z : Calcd for C₄₇H₅₉NNaO₁₀Si [M+Na]⁺: 848.4, found: 848.4; Elemental analysis: Calcd (%) for C₄₇H₅₉NO₁₀Si: C 68.34, H 7.20, N 1.70, found: C 68.46, H 7.27, N 1.66.

2-(Trimethylsilyl)ethyl (2,3,4-tri-*O*-benzyl- α -L-fucopyranosyl)-(1 \rightarrow 3)-2-acetamido-6-*O*-benzyl-2-deoxy- β -D-glucopyranoside (S4**).** To a solution of **S3** (600 mg, 0.73 mmol) in anhydrous THF (14 mL) were added Me₃N·BH₃ (212 mg, 2.91 mmol) and anhydrous AlCl₃ (580 mg, 4.35 mmol) with stirring under argon at r.t. When all reagents were dissolved, H₂O (26 μ L) was added and the mixture was stirred for 5 h. The reaction was quenched by addition of H₂O (12 mL) and 1 N aq HCl (12 mL) and extracted with CH₂Cl₂ (3 \times 30 mL). The organic layers were washed with brine (30 mL), dried over Na₂SO₄, and concentrated. Column chromatography on silica gel (petroleum ether/EtOAc, 7:3 to 1:1) afforded **S4** (470 mg, 0.57 mmol, 78%) as white foam.

¹H NMR (500.1 MHz, CDCl₃): δ = 7.43-7.23 (m, 20H, Ar-H), 5.51 (d, J = 7.1 Hz, 1H, NH), 4.96 (m, 2H, Fuc-H1, CH₂Ph), 4.91 (d, J = 8.2 Hz, GlcNAc-H1), 4.86-4.57 (m, 7H, CH₂Ph), 4.13-4.04 (m, 2H, Fuc-H2, Fuc-H5), 4.02-3.89 (m, 3H, Fuc-H3, GlcNAc-H3, CH₂CH₂Si), 3.83 (m, 1H, GlcNAc-H6a), 3.73-3.66 (m, 2H, Fuc-H4, GlcNAc-H6b), 3.59-3.47 (m, 2H, GlcNAc-H5, CH₂CH₂Si), 3.43 (dd, J = 8.8 Hz, 1H, GlcNAc-H4), 3.26 (m, GlcNAc-H2), 1.60 (s, 3H, COCH₃), 1.15 (d, J = 6.3 Hz, Fuc-H6), 1.02-0.84 (m, 2H, CH₂CH₂Si), -0.01 (s, 9H,

Si(CH₃)₃); ¹³C NMR (125.8 MHz, CDCl₃): δ = 170.9 (COCH₃), 138.6, 138.4, 128.7, 128.6, 128.5, 128.4, 128.1, 128.0, 127.9, 127.8, 127.7, 127.6 (24C, Ar-C), 99.4, 99.3 (Fuc-C1, GlcNAc-C1), 84.1 (GlcNAc-C3), 79.3 (Fuc-C3), 77.4 (Fuc-C4), 76.0 (Fuc-C2), 75.1 (CH₂Ph), 75.0 (GlcNAc-C5), 74.1, 73.6, 73.0 (CH₂Ph), 70.7 (GlcNAc-C4), 69.8 (GlcNAc-C6), 68.1 (Fuc-C5), 67.0 (CH₂CH₂Si), 56.7 (GlcNAc-C2), 23.4 (COCH₃), 18.2 (CH₂CH₂Si), 16.8 (Fuc-C6), -1.3 (3C, Si(CH₃)₃); [α]_D -36.7 (c 0.55, CHCl₃); MS (ESI): *m/z*: Calcd for C₄₇H₆₁NNaO₁₀Si [M+Na]⁺: 850.4, found: 850.5.

2-(Trimethylsilyl)ethyl (benzyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosynate)-(2→3)-2,4-di-*O*-acetyl-6-*O*-benzoyl-β-D-galactopyranosyl-(1→4)-[2,3,4-tri-*O*-benzyl-α-L-fucopyranosyl-(1→3)]-2-acetamido-6-*O*-benzyl-2-deoxy-β-D-glucopyranoside (S6). Compounds **S5** (300 mg, 0.31 mmol) and **S4** (380 mg, 0.46 mmol) were dissolved in anhydrous CH₂Cl₂ (8 mL). Powdered activated molecular sieves 4 Å (0.8 g) were added and the mixture was stirred at r.t. under argon. After 3.5 h, a solution of DMTST (200 mg, 0.77 mmol) in anhydrous CH₂Cl₂ (2.0 mL), which had been stirred with molecular sieves 4 Å (0.2 g) for 3.5 h, was added. After stirring for 5 d, the solution was diluted with CH₂Cl₂ (10 mL), filtered and successively washed with satd aq NaHCO₃ (20 mL) and brine (20 mL). The aqueous layers were extracted with CH₂Cl₂ (3 × 20 mL) and the combined organic layers were dried over Na₂SO₄ and concentrated under reduced pressure. Column chromatography on silica (petroleum ether/EtOAc/MeOH, 8:5:0.5 to 8:5:0.9) afforded **S6** (110 mg, 0.06 mmol, 20%) as white solid.

¹H NMR (500.1 MHz, CDCl₃): δ = 8.09-7.21 (m, 30H, Ar-H), 6.22 (d, *J* = 8.1 Hz, 1H, GlcNAc-NH), 5.62 (ddd, *J* = 2.5, 5.9, 8.7 Hz, 1H, Sia-H8), 5.43-5.32 (m, 2H, Sia-H7, PhCH₂), 5.22 (d, *J* = 3.5 Hz, 1H, Fuc-H1), 5.20 (d, *J* = 3.5 Hz, 1H, Gal-H4), 5.07-4.95 (m, 3H, Gal-H2, PhCH₂), 4.92 (m, 1H, Sia-H4), 4.88-4.72 (m, 7H, Gal-H1, Gal-H3, GlcNAc-H1, PhCH₂), 4.68 (B of AB, *J* = 11.8 Hz, 1H, PhCH₂), 4.61 (A of AB, *J* = 12.0 Hz, 1H, PhCH₂), 4.61 (B of AB, *J* = 12.0 Hz, 1H, PhCH₂), 4.41-4.29 (m, 2H, Gal-H6a, Sia-H9a), 4.23 (dd, *J* = 7.0, 11.0 Hz, 1H, Gal-H6b), 4.20-3.96 (m, 8H, Fuc-H2, Fuc-H5, Gal-H5, GlcNAc-H3, GlcNAc-H4, GlcNAc-H6a, Sia-H5, Sia-H9b), 3.93-3.78 (m, 5H, Fuc-H3, GlcNAc-H2, GlcNAc-H5, GlcNAc-H6b, CH₂CH₂Si), 3.56-3.47 (m, 3H, Fuc-H4, Sia-H6, CH₂CH₂Si), 2.64 (dd, *J* = 4.5, 12.6 Hz, 1H, Sia-H3^{eq}), 2.27, 2.15, 2.11, 2.03, 2.00, 1.94, 1.88, (7s, 24H, 8 COCH₃), 1.75 (t, *J* = 12.4 Hz, 1H, Sia-H3^{ax}), 1.13 (d, *J* = 6.5 Hz, 3H, Fuc-H6), 1.06-0.77 (m, 2H, CH₂CH₂Si), 0.00 (s, 9H, Si(CH₃)₃); ¹³C NMR (125.8 MHz, CDCl₃): δ = 170.7, 170.6,

170.5, 170.4, 170.1, 169.8 (8C, COCH₃), 167.5 (Sia-C1), 165.8 (ArCO), 139.2, 139.1, 138.8, 138.6, 134.8, 133.4, 129.8, 129.0, 128.8, 128.6, 128.5, 128.4, 128.3, 127.8, 127.7, 127.6, 127.5, 127.3 (36C, Ar-C), 99.2, 99.1 (Gal-C1, GlcNAc-C1), 97.0 (Sia-C2), 96.5 (Fuc-C1), 79.5 (Fuc-C3), 77.7 (Fuc-C4), 76.6 (Fuc-C2), 74.6 (PhCH₂), 74.2, 73.5, 73.4 (GlcNAc-C3, GlcNAc-C4, GlcNAc-C5), 73.0, 72.9 (3C, 3 CH₂Ph), 72.2 (Sia-C6), 71.2, 70.7, 70.4, (Gal-C2, Gal-C3, Gal-C5), 69.5 (GlcNAc-C6), 69.4 (Sia-C4), 68.5 (PhCH₂), 67.8, 67.7, 67.2 66.9 (Fuc-C5, Gal-C4, Sia-C7, Sia-C8), 66.6 (CH₂CH₂Si), 62.6 (Sia-C9), 61.7 (Gal-C6), 53.4 (GlcNAc-C2), 49.0 (Sia-C5), 37.6 (Sia-C3), 23.3, 23.2, 21.5, 21.0, 20.9, 20.8, 20.7 (8C, 8 COCH₃), 18.0 (CH₂CH₂Si), 16.8 (Fuc-C6), -1.3 (3C, Si(CH₃)₃); [α]_D -24.3 (*c* 0.58, CHCl₃); MS (ESI): *m/z*: Calcd for C₉₀H₁₁₀N₂NaO₃₀Si [M+Na]⁺: 1749.7, found: 1749.8.

2-(Trimethylsilyl)ethyl (sodium 5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosynate)-(2→3)-β-D-galactopyranosyl-(1→4)-[α-L-fucopyranosyl-(1→3)]-2-acetamido-2-deoxy-β-D-glucopyranoside (sLe^x-OTMSE). Compound **S6** (90 mg, 0.052 mmol) was dissolved in dioxane/water (4:1, 10 mL) under argon. Pd(OH)₂/C (20 mg, 10% Pd) was added and the resulting mixture was hydrogenated (4 bar H₂) at r.t. After 24 h, the mixture was filtered and the solvent was removed under reduced pressure. The residue was redissolved in 1 N aq NaOH (10 mL) and MeOH (2 mL) and stirred at r.t. After 20 h the mixture was neutralized with 1 N HCl, the volatiles were removed under reduced pressure, and the crude product was purified by size-exclusion chromatography and RP chromatography (RP18, H₂O/MeOH). Lyophilization from water afforded **sLe^x-OTMSE** (21 mg, 0.022 mmol, 43%) as white fluffy solid.

¹H NMR (500.1 MHz, D₂O/CD₃OD): δ = 5.09 (d, *J* = 4.0 Hz, 1H, FucH-1), 4.80 (m, FucH-5), 4.56 (d, *J* = 8.1 Hz, 1H, GlcNAc-H1), 4.52 (d, *J* = 7.8 Hz, 1H, Gal-H1), 4.08 (dd, *J* = 3.1, 9.8 Hz, 1H, Gal-H3), 4.06-3.55 (m, 22H, CH₂CH₂Si, Fuc-H2, Fuc-H3, Fuc-H4, Gal-H4, Gal-H5, Gal-H6a, Gal-H6b, GlcNAc-H2, GlcNAc-H3, GlcNAc-H4, GlcNAc-H5, GlcNAc-H6a, GlcNAc-H6b, Sia-H4, Sia-H5, Sia-H6, Sia-H7, Sia-H8, Sia-H9a, Sia-H9b), 3.52 (dd, *J* = 7.8, 9.8 Hz, Gal-H2), 2.76 (dd, *J* = 4.6, 12.1 Hz, 1H, Sia-H3^{eq}), 2.03, 2.01 (2s, 6H, 2 COCH₃), 1.79 (t, *J* = 12.1 Hz, 1H, Sia-H3^{ax}), 1.17 (d, *J* = 6.6 Hz, 3H, Fuc-H6), 1.05-0.80 (m, 2H, CH₂CH₂Si), 0.00 (s, 9H, Si(CH₃)₃); ¹³C NMR (125.8 MHz, D₂O/CD₃OD): δ = 176.0, 175.1, 174.8 (2 NHCOCH₃, Sia-C1) 102.6 (Gal-C1), 101.2, 100.6 (GlcNAc-C1, Sia-C2), 100.6 (Fuc-C1), 99.6, 76.6, 76.3, 76.0, 75.9, 74.4, 73.9, 72.9, 72.8, 70.2, 69.3, 69.1, 68.7, 68.3, 67.7, 63.6, 62.4, 60.6, 56.7, 52.7, 40.8 (Sia-C3), 23.3, 23.0 (2 NHCOCH₃), 18.1

(CH₂CH₂Si), 16.2 (Fuc-C6), -1.5 (3C, Si(CH₃)₃); [α]_D -43.9 (c 0.70, MeOH); HRMS (ESI): *m/z*: Calcd for C₃₆H₆₄N₂NaO₂₃Si⁺ [M+H]⁺: 943.3561, found: 943.3553; HPLC purity: > 99.5%.

Molecular cloning

Human E-selectin cDNA encoding the native E-selectin secretion signal, the Lec-, EGF- and first two SCR domains (E-selectin*, residues 1–280) was isolated and PCR amplified from previously generated full length human E-selectin-IgG(Fc) expressing Chinese hamster ovary (CHO-K1) cells.⁸ The Kozak consensus sequence was introduced to increase protein yield. E-selectin* cDNA was cloned into the pcDNA3.1(+) expression vector (Life Technologies Ltd, UK) using standard cloning procedures and sequence verified.

Stably transfected CHO-K1 cells (ATCC-No. CCL-61) were generated using FuGENE[®] HD Transfection Reagent (Roche Applied Science, Switzerland), G418 treatment, and subsequent clone selection. Single clones were screened for optimal protein expression. The CHO-K1 cells were cultivated as monolayer in Ham's Nutrient Mixture F-12 medium supplemented with 2 mM L-glutamine (Invitrogen, Lucerne, Switzerland), 5-10% (v/v) fetal calf serum (Invitrogen, Lucerne, Switzerland), 100 U/mL penicillin, and 100 µg/mL streptomycin. Once confluent, the transfected CHO-K1 cells were treated with culture medium containing 1 µM kifunensine (Cayman Chemical, MI) for 2 days. The medium was discarded and replaced by culture medium containing 5 µM kifunensine. Conditioned culture medium was collected weekly.

The following primers were used for amplification of E-selectin* cDNA:

SELE_koz_fw with *EcoRI* restriction site and Kozak consensus sequence

5'–GCGAATTCGCCGCCACCATGGGAATTGCTTCACAGTTTCTCTCAG–3'

SELE_SCR2_rv with two stop codons and an *EcoRV* restriction site

5'–GCGATATCCTATTAAGCTTTACACGTTGGCTTCTC–3'

Protein purification

A purification column with the functional anti-E-selectin antibody 7A9 coupled to sepharose was prepared as previously described.⁹ Affinity chromatography was performed with a functional anti-E-selectin antibody (7A9) column with loading buffer (50 mM Tris/HCl, pH 7.6, 150 mM NaCl, and 0.05% v/v Tween-20), washing buffer (5 mM NH₄OAc/CH₃COOH, pH 5.0), and elution buffer (0.5 M CH₃COOH/NH₄OAc, pH 3.4). Protein samples were adjusted to neutral pH upon elution with 2.5 M Tris and concentrated by ultrafiltration (10

kDa cut-off, Millipore, MA). Approximately 5 mg high-mannosylated E-selectin was obtained per liter medium.

Concentrated high-mannosylated E-selectin (10-20 mg/mL) was deglycosylated with endoglycosidase H_f (New England Biolabs, MA) according to the recommendation of the manufacturer using 10 U/μg protein for 3 h at 37 °C under non-reducing conditions. The deglycosylation yielded protein with a single GlcNAc moiety attached to each *N*-linked glycosylation site. The deglycosylation progress was checked with standard SDS-PAGE/Coomassie analysis. When required, deglycosylation was extended with additional 5 U EndoH_f per μg protein for 1 h. Deglycosylated protein was re-purified over the functional anti-E-selectin antibody (7A9) column and concentrated to 50 mg/mL in 10 mM Tris-HCl, pH 7.4, 5 mM CaCl₂.

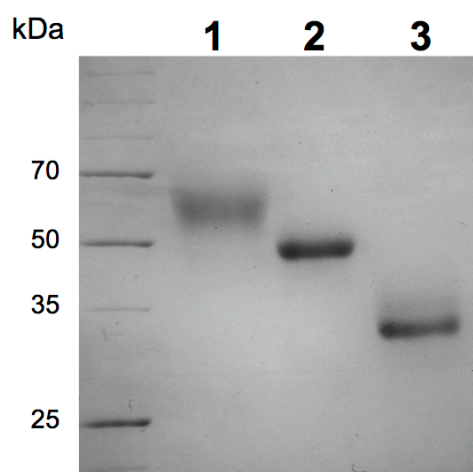
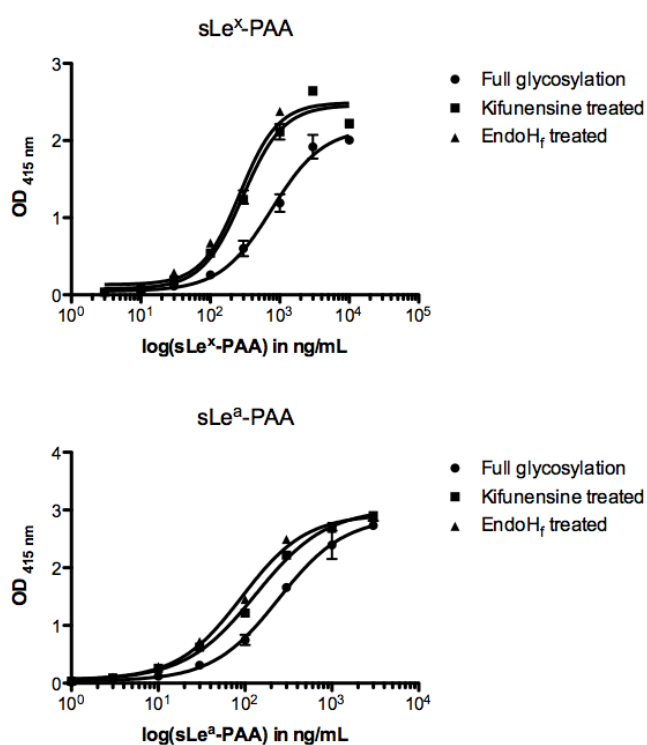


Figure S1. SDS-PAGE (12%) of E-selectin* expressed CHO cells in absence (lane 1) or presence (lane 2) of kifunensine, and endoglycosidase H treated protein (lane 3).

Activity assays

To verify the activity of E-selectin upon deglycosylation, the affinity towards biotinylated-polyacrylamide (PAA) functionalized with sLe^x (Neu5Ac(α2-3)Gal(β1-4)[Fuc(α1-3)]GlcNAc) and its isoform sLe^a (Neu5Ac(α2-3)Gal(β1-3)[Fuc(α1-4)]GlcNAc) was determined similarly as previously described.⁹ Briefly, fully glycosylated E-selectin expressed in absence of kifunensine, high-mannosylated E-selectin, and endoglycosidase H treated E-selectin were coated at a concentration of 0.2 μM on polystyrene dishes (Immunoplate II MaxiSorp™ 96-well microtiter plates, Nunc, Denmark) at 4 °C overnight in assay buffer (10 mM Tris/HCl, pH 7.4, 150 mM NaCl, 1 mM CaCl₂) and blocked with

bovine serum albumin (2% w/v in assay buffer). After washing with assay buffer, serial dilutions of sLe^x-PAA or sLe^a-PAA were applied, which were co-incubated with streptavidin-peroxidase (200 U/mL, Roche Applied Science, Rotkreuz, Switzerland) prior to the assay. Control wells without protein were treated in the same fashion. Carbohydrate-PAA binding was detected by colorimetric reaction with 2,2'-azino-di-(3-ethylbenzthiazoline)-6-sulfonic acid (Bio-Rad, Reinach BL, Switzerland), which was stopped with 2% (w/v) oxalic acid. The optical density was measured at 415 nm. Measured values were subtracted with the control wells and curve fitting was performed with the Prism software (GraphPad Software Inc., CA). A four-parameter fit for minimum, maximum, logEC₅₀ and slope was used. The EC₅₀ values correspond to the half-maximal effective concentration (Figure S2).



	EC ₅₀ (ng/mL) sLe ^x -PAA	EC ₅₀ (ng/mL) sLe ^a -PAA
Fully glycosylated	762	243
Kifunensine treated	292	132
EndoH	267	92

Figure S2. Activity assays as described above of fully glycosylated, high-mannosylated, and deglycosylated E-selectin* with sLe^x-polyacrylamide (PAA) and sLe^a-PAA.

Protein-ligand co-crystallization

Deglycosylated E-selectin* was co-crystallized with sLe^x-OTMSE or glycomimetic **1** in a sitting drop vapor diffusion setup. Initial crystals were obtained in equal volumes of 23.5% (w/v) polyethylene glycol 8000, 0.1 M 3-(*N*-morpholino)propanesulfonic acid pH 6.2, 0.2 M CaCl₂ with protein (25 mg/mL) containing each ligand (final concentration 25 mM sLe^x or 10 mM glycomimetic **1**). Diffraction quality crystals were obtained in 12.5-13.5 % PEG8000, MOPS 0.1 M pH 6.2, 0.2 M CaCl₂ after microseeding within 48 h. Plate-like crystals appeared within one day. The drops were covered with perfluoropolyether (Hampton Research, CA) prior to flash cooling in liquid nitrogen. Diffraction data were collected at the PXI beamline at the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland. Data were processed with the XDS software.¹⁰

Structure determination and refinement

The structure of E-selectin*/**1** was determined by molecular replacement using PHASER¹¹ within the CCP4 software package¹² with the P-selectin/PSGL-1 (PDB code 1G1S)¹³ structure for the Lec- and EGF domains, human mannose-binding lectin-associated serine protease-1 (PDB code 3GOV)¹⁴ for the first SCR, and human CD55 (PDB code 1H04)¹⁵ for the second SCR. The E-selectin*/sLe^x structure was determined with the E-selectin*/**1** structure. Refinement was performed with the BUSTER¹⁶ software and model building was performed with COOT.¹⁷ Restraints for the ligands were generated using PRODRG.¹⁸ Modeling of the TMSE protecting group of the sLe^x ligand was omitted due to poor density of this flexible group. Parts of the SCR2 domain in the E-selectin*/sLe^x structure were modeled on the basis of the E-selectin*/**1** structure due to weak density. In the E-selectin*/sLe^x structure, 15 residues are within the allowed Ramachandran region (2.7%) and 11 are outliers (2.0%). In the E-selectin*/**1** structure, 14 residues are within the allowed region (2.9%) and 7 are outliers (1.4%). Data collection and refinement statistics are listed in Table S1. Root-mean-square deviations (rmsd) for the Lec- and EGF domains were calculated with the DALI server.¹⁹ Figures were generated with the Visual Molecular Dynamics (VMD) software.²⁰

Table S1. Data collection and refinement statistics.

	E-selectin/sLe ^x	E-selectin/1
PDB Entry	4CSY	4C16
Data collection		
Wavelength (Å)	0.99985	0.99987
Space group	<i>P1</i>	<i>P1</i>
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	52.3, 58.7, 58.9	52.6, 56.9, 59.7
α , β , γ (°)	76.1, 86.3, 86.3	74.4, 87.1, 88.3
Molecules in the asymmetric unit	2	2
Resolution (Å)*	28.3-2.40 (2.53-2.40)	57.4-1.93 (2.03-1.93)
<i>R</i> _{merge} *	0.043 (0.317)	0.040 (0.257)
<i>I</i> / σI *	8.8 (2.0)	9.3 (2.0)
Completeness (%)*	95.4 (90.2)	92.5 (91.0)
Redundancy*	1.8 (1.8)	1.8 (1.7)
Refinement		
Resolution (Å)*	28.3-2.40 (2.50-2.40)	15.2-1.93 (1.98-1.93)
No. reflections*	25130 (2538)	46243 (3357)
<i>R</i> _{work} / <i>R</i> _{free} %	21.7 / 25.3	18.3 / 23.2
No. atoms		
Protein	4387	4571
Ligand/ion	312	280
Water	127	525
<i>B</i> -factors		
Protein	78.2	39.7
Ligand/ion	101.1	52.8
Water	63.2	48.3
RMS deviations		
Bond lengths (Å)	0.010	0.010
Bond angles (°)	1.22	1.16

RMS indicates root mean square. Single crystals were used.

*Values in parentheses are for highest-resolution shell.

Small angle X-ray scattering

Small-angle X-ray scattering (SAXS) measurements were conducted at the X12SA-cSAXS beamline (Swiss Light Source) with a sample to detector (Pilatus 2M) distance of 7.0 m and ten measurements of 0.5 s exposure time. Deglycosylated E-selectin* was diluted to 10 mg/mL and measured either in presence or absence of 3 mM glycomimetic **1**. Buffer samples (10 mM Tris/HCl, pH 7.4, 5 mM CaCl₂) alone were measured as reference. A control

measurement with 3 mM glycomimetic **1** without protein showed no difference to the buffer control. Integrated data were background subtracted with the PRIMUS software.²¹ Experimental and theoretical scattering curves were evaluated with the CRY SOL software.²² A model for the E-selectin* bent conformation was generated using the coordinates of the Lec-EGF domains of the apo-E-selectin (PDB code 1ESL)²³ structure combined with the coordinates of the two SCR domains from E-selectin*/**1**. Data evaluation and fitting parameters are given in Table S2.

Table S2. SAXS parameters for data validation and interpretation.

	E-selectin, ligand absent		E-selectin, ligand present	
Fit Model	E-selectin bent	E-selectin extended	E-selectin bent	E-selectin extended
Goodness of fit (χ^2)	0.622	0.702	0.891	0.758
q-range	0.007 - 0.500 Å ⁻¹	0.007 - 0.500 Å ⁻¹	0.007 - 0.500 Å ⁻¹	0.007 - 0.500 Å ⁻¹
R _g (guinier plot)	37.84 Å	37.84 Å	41.80 Å	41.80 Å
R _g model	40.07 Å	41.02 Å	40.07 Å	41.02 Å
D _{max} experimental	132.43 Å	132.43 Å	147.56 Å	147.56 Å
D _{max} model	137.23 Å	143.58 Å	137.23 Å	143.58 Å

$$q = 4\pi\sin(\theta)/\lambda$$

D_{max} indicates the maximum particle size.

R_g indicates the radius of gyration as calculated with the AutoRG software using globular fitting parameters.

Molecular dynamics simulations

The Protein Preparation Wizard²⁴ was used to convert the extended and bent (PDB code 1G1T)¹³ E-selectin crystal structures to all-atom protein models. A rigid receptor grid was generated from the crystal structure 1G1T, using the OPLS 2005 force field.^{25,26} An initial protein-ligand complex of glycomimetic **1** docked to 1G1T was obtained through flexible docking using Glide.²⁷ The best docking pose was selected interactively considering the proposed binding mode already known from existing structures and similar ligands. Molecular dynamics simulations were carried out with Desmond²⁸ using the OPLS 2005 force field, as implemented in the Schrödinger 2012 suite.²⁴ Each system was solvated using an orthorhombic, TIP3P water box²⁹ with a minimum distance of 10 Å from the complex. Counterions (Na⁺ and Cl⁻) were added to neutralize the charges and account for physiological salt concentration (0.15 M). Long-range electrostatic interactions were handled using the particle mesh Ewald summation.³⁰ All systems were equilibrated using the default relaxation protocol and simulated over the span of 48 ns with a time step of 2.0 fs. The SHAKE algorithm³¹ was applied to all heavy atom-bound hydrogens. Production runs were carried out

in the Martyna-Tobias-Klein isothermal-isobaric ensemble (NPT)³² using the Nose-Hoover barostat to maintain a constant temperature of 300 K.³³ Energetic and structural data were recorded in 4.8 ps intervals. To assess the stability of the MD simulations, energetic and structural data was analyzed over the course of the entire simulation time (Table S3).

Table S3. Energetic and structural values of the MD simulations.

Conformation	Ligand	Average total energy [kcal/mol]	Slope* [ps ⁻¹]	Average backbone RMSD
Bent	sLe ^x	-68804 ± 250	0.018	0.473 ± 0.019
Bent	Glycomimetic 1	-70675 ± 215	0.015	0.481 ± 0.019
Extended	sLe ^x	-69939 ± 367	0.026	0.486 ± 0.019
Extended	Glycomimetic 1	-70557 ± 289	0.020	0.507 ± 0.017

* The slope denotes the change of the total energy over time.

From the collected MD frames, every tenth frame (1000 frames in total) was used to calculate the average ΔG applying the molecular mechanics generalized Born surface area (MM-GB/SA) method implemented in Prime³⁴ (Table S4).

Table S4. Calculated interaction energies of sLe^x and glycomimetic **1** to both E-selectin conformations.

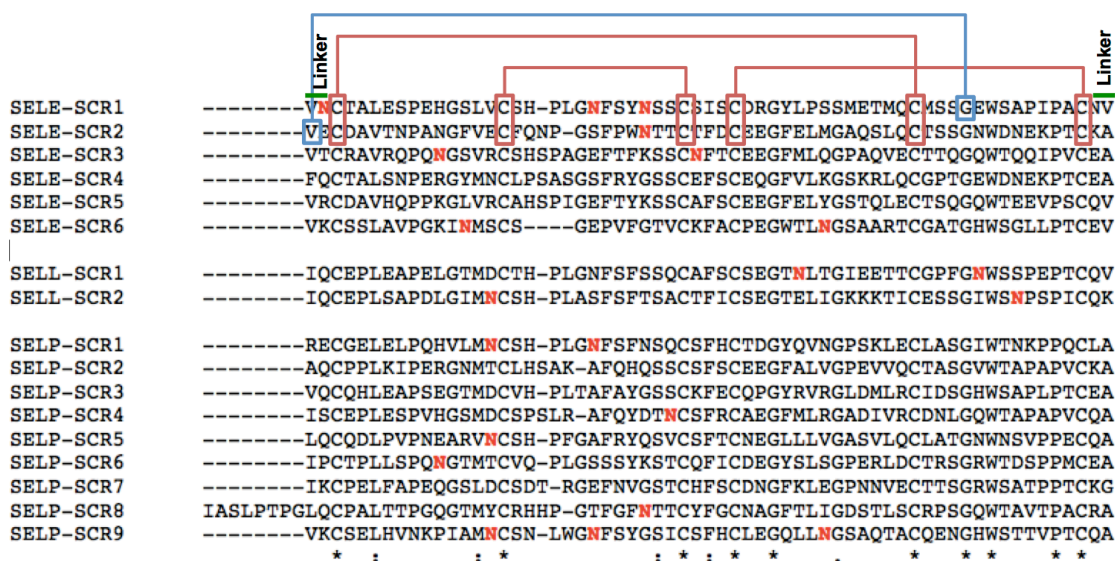
Conformation	Ligand	Average ΔG [kcal/mol]	Scaled ΔG [kcal/mol]
Bent	sLe ^x	-45.1 ± 5.2	-3.36
Bent	Glycomimetic 1	-38.3 ± 5.0	-2.86
Extended	sLe ^x	-56.1 ± 6.0	-4.18
Extended	Glycomimetic 1	-58.7 ± 6.8	-4.38

The observed high ΔG energies as well as the high dynamic range in comparison to published experimental data⁹ is due to limitations of the MM-GB/SA method. Protein desolvation energies, enthalpy–entropy compensation, or overestimation of electrostatic interactions as described by Guimarães *et al.*^{35,36} are causes of energy overestimation with this method. For a more realistic representation of the binding energies, we scaled the calculated results by a factor of 13.4 based on published ITC measurements of sLe^x.⁹

SCR comparison

Table S5. List of SCR domain structures solved by X-ray crystallography deposited on the PDB.

PDB Entry	No. of SCR domains	No. of disulfide bonds per SCR	Ref
This publication	2	3	-
1C1Z	4	2	37
1QUB			38
1CKL	2	2	39
1H2P (= 1H2Q, 1H03, 1H04, 1UOT)	2	2	15
1LY2	2	2	40
1OJV (= 1OJW, 1OJY, 1OK1, 1OK2, 1OK3, 1OK9)	3	2	41
1ZJK	2	2	42
2W80 (= 2W81)	1	2	43
2XRD	3	2	44
2UWN (= 2V8E)	2	2	45

**Figure S3.** Alignment of E- (SELE), L- (SELL), and P-selectin (SELP) short consensus repeat domains (SCR). The disulfide bonds are connected in red. The stabilizing loop glycine interaction with the linker region is marked in blue. Residues marked red are potential N-linked glycosylation sites. Sequence alignment was performed with the ClustalW2 server.⁴⁶

Lec-EGF comparison



Figure S4. Sequence alignment of E- and P-selectin Lec (residues 1-120) and EGF-like (residues 121-157) domains. Underlined residues are involved in ligand binding and/or Ca^{2+} coordination in the extended conformation (taken from PDB code 1G1S for P-selectin)¹³. Residues marked red are potential *N*-linked glycosylation sites.

Supplementary References

- (1) Gottlieb, H. E.; Kotlyar, V.; Nudelman, A. *J. Org. Chem.* **1997**, *62*, 7512.
- (2) Thoma, G.; Magnani, J. L.; Ohrlein, R.; Ernst, B.; Schwarzenbach, F.; Duthaler, R. O. *J. Am. Chem. Soc.* **1997**, *119*, 7414.
- (3) Takaku, H.; Sato, J.; Ishida, H.; Inazu, T.; Ishida, H.; Kiso, M. *Glycoconj. J.* **2006**, *23*, 565.
- (4) Lemieux, R. U.; Hendriks, K. B.; Stick, R. V.; James, K. *J. Am. Chem. Soc.* **1975**, *97*, 4056.
- (5) Sherman, A. A.; Mironov, Y. V.; Yudina, O. N.; Nifantiev, N. E. *Carbohydr. Res.* **2003**, *338*, 697.
- (6) Fugedi, P.; Garegg, P. J. *Carbohydr. Res.* **1986**, *149*, C9.
- (7) Bhunia, A.; Schwardt, O.; Gathje, H.; Gao, G. P.; Kelm, S.; Benie, A. J.; Hricovini, M.; Peters, T.; Ernst, B. *Chembiochem* **2008**, *9*, 2941.
- (8) Scheffler, K.; Ernst, B.; Katopodis, A.; Magnani, J. L.; Wang, W. T.; Weisemann, R.; Peters, T. *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1841.
- (9) Binder, F. P.; Lemme, K.; Preston, R. C.; Ernst, B. *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 7327.
- (10) Kabsch, W. *J Appl Crystallogr* **1993**, *26*, 795.
- (11) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J Appl Crystallogr* **2007**, *40*, 658.
- (12) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. *Acta Crystallogr D* **2011**, *67*, 235.
- (13) Somers, W. S.; Tang, J.; Shaw, G. D.; Camphausen, R. T. *Cell* **2000**, *103*, 467.
- (14) Dobo, J.; Harmat, V.; Beinrohr, L.; Sebestyen, E.; Zavodszky, P.; Gal, P. *J. Immunol.* **2009**, *183*, 1207.
- (15) Williams, P.; Chaudhry, Y.; Goodfellow, I. G.; Billington, J.; Powell, R.; Spiller, O. B.; Evans, D. J.; Lea, S. *J. Biol. Chem.* **2003**, *278*, 10691.
- (16) Blanc, E.; Roversi, P.; Vonnrhein, C.; Flensburg, C.; Lea, S. M.; Bricogne, G. *Acta Crystallogr D* **2004**, *60*, 2210.
- (17) Emsley, P.; Cowtan, K. *Acta Crystallogr D* **2004**, *60*, 2126.
- (18) Schuttelkopf, A. W.; van Aalten, D. M. F. *Acta Crystallogr D* **2004**, *60*, 1355.
- (19) Holm, L.; Kaariainen, S.; Rosenstrom, P.; Schenkel, A. *Bioinformatics* **2008**, *24*, 2780.
- (20) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graph. Model.* **1996**, *14*, 33.
- (21) Konarev, P. V.; Volkov, V. V.; Sokolova, A. V.; Koch, M. H. J.; Svergun, D. I. *J Appl Crystallogr* **2003**, *36*, 1277.
- (22) Svergun, D.; Barberato, C.; Koch, M. H. J. *J Appl Crystallogr* **1995**, *28*, 768.
- (23) Graves, B. J.; Crowther, R. L.; Chandran, C.; Rumberger, J. M.; Li, S.; Huang, K. S.; Presky, D. H.; Familletti, P. C.; Wolitzky, B. A.; Burns, D. K. *Nature* **1994**, *367*, 532.
- (24) Schrödinger; Maestro; version 9.3; LLC: New York, NY, **2012**.
- (25) Jorgensen, W. L.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1988**, *110*, 1657.
- (26) Jorgensen, W. L.; Maxwell, D. S.; TiradoRives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225.
- (27) Schrödinger; Glide; version 5.8; LLC: New York, NY, **2012**.
- (28) Desmond Molecular Dynamics System; version 3.1; D.E. Shaw Research: New York, NY, **2012**.

- (29) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926.
- (30) Darden, T.; York, D.; Pedersen, L. *J. Chem. Phys.* **1993**, *98*, 10089.
- (31) Ryckaert, J. P.; Ciccotti, G.; Berendsen, H. J. C. *J Comput Phys* **1977**, *23*, 327.
- (32) Martyna, G. J.; Tobias, D. J.; Klein, M. L. *J. Chem. Phys.* **1994**, *101*, 4177.
- (33) Nose, S. *Mol Phys* **1984**, *52*, 255.
- (34) Schrödinger; Prime; version 3.1; LLC: New York, NY, **2012**.
- (35) Guimaraes, C. R.; Cardozo, M. *J. Chem. Inf. Model.* **2008**, *48*, 958.
- (36) Guimaraes, C. R.; Mathiowetz, A. M. *J. Chem. Inf. Model.* **2010**, *50*, 547.
- (37) Schwarzenbacher, R.; Zeth, K.; Diederichs, K.; Gries, A.; Kostner, G. M.; Laggner, P.; Prassl, R. *EMBO J.* **1999**, *18*, 6228.
- (38) Bouma, B.; de Groot, P. G.; van den Elsen, J. M. H.; Ravelli, R. B. G.; Schouten, A.; Simmelink, M. J. A.; Derksen, R. H. W. M.; Kroon, J.; Gros, P. *EMBO J.* **1999**, *18*, 5166.
- (39) Casanovas, J. M.; Larvie, M.; Stehle, T. *EMBO J.* **1999**, *18*, 2911.
- (40) Prota, A. E.; Sage, D. R.; Stehle, T.; Fingerioth, J. D. *Proc. Natl. Acad. Sci. U. S. A.* **2002**, *99*, 10641.
- (41) Lukacik, P.; Roversi, P.; White, J.; Esser, D.; Smith, G. P.; Billington, J.; Williams, P. A.; Rudd, P. M.; Wormald, M. R.; Harvey, D. J.; Crispin, M. D. M.; Radcliffe, C. M.; Dwek, R. A.; Evans, D. J.; Morgan, B. P.; Smith, R. A. G.; Lea, S. M. *Proc. Natl. Acad. Sci. U. S. A.* **2004**, *101*, 1279.
- (42) Gal, P.; Harmat, V.; Kocsis, A.; Bian, T.; Barna, L.; Ambrus, G.; Vegh, B.; Balczer, J.; Sim, R. B.; Naray-Szabo, G.; Zavodszky, P. *J. Biol. Chem.* **2005**, *280*, 33435.
- (43) Schneider, M. C.; Prosser, B. E.; Caesar, J. J.; Kugelberg, E.; Li, S.; Zhang, Q.; Quoraishi, S.; Lovett, J. E.; Deane, J. E.; Sim, R. B.; Roversi, P.; Johnson, S.; Tang, C. M.; Lea, S. M. *Nature* **2009**, *458*, 890.
- (44) Roversi, P.; Johnson, S.; Caesar, J. J. E.; McLean, F.; Leath, K. J.; Tsiftoglou, S. A.; Morgan, B. P.; Harris, C. L.; Sim, R. B.; Lea, S. M. *Acta Crystallogr F* **2011**, *67*, 739.
- (45) Prosser, B. E.; Johnson, S.; Roversi, P.; Herbert, A. P.; Blaum, B. S.; Tyrrell, J.; Jowitt, T. A.; Clark, S. J.; Tarelli, E.; Uhrin, D.; Barlow, P. N.; Sim, R. B.; Day, A. J.; Lea, S. M. *J. Exp. Med.* **2007**, *204*, 2277.
- (46) Kim, T.; Joo, H. *BMC Bioinformatics* **2010**, *11*.

Chapter 2.4. – Manuscript 2

Acid Pre-Organization of a Glycomimetic E-selectin Antagonist

Manuscript in preparation. Formatted according to the requirements of *Angewandte Int. Ed.*

Contributions

- | | |
|---------------|--|
| R. C. Preston | <ul style="list-style-type: none">• Protein preparation.• Protein crystallography.• Data collection, processing, and model building.• Manuscript preparation. |
| K. Lemme | <ul style="list-style-type: none">• Isothermal titration calorimetry. |
| C. Weckerle | <ul style="list-style-type: none">• Surface plasmon resonance. |
| M. Smieško | <ul style="list-style-type: none">• Molecular dynamics simulations. |
| R. P. Jakob | <ul style="list-style-type: none">• Protein crystallization support. |
| T. Maier | <ul style="list-style-type: none">• Support in data collection, processing, and model building. |
-

Acid Pre-Organization of a Glycomimetic E-selectin Antagonist**

Roland C. Preston¹, Katrin Lemme¹, Céline Weckerle¹, Martin Smieško¹, Roman Jakob², Timm Maier², Beat Ernst^{1*}

¹ Institute of Molecular Pharmacy, Pharmacenter, University of Basel, 4056 Basel, Switzerland

² Department of Biochemistry, Biocenter, University of Basel, 4056 Basel, Switzerland

[*] R. C. Preston, Dr. Céline Weckerle, Dr. Katrin Lemme, Dr. Martin Smieško, Prof. Dr. B. Ernst

Institute of Molecular Pharmacy, University of Basel

Klingelbergstrasse 50, 4056 Basel (Switzerland)

E-mail: beat.ernst@unibas.ch

Dr. R. Jakob, Prof. Dr. T. Maier

Biocenter, University of Basel

Klingelbergstrasse 70, 4056 Basel (Switzerland)

[**] The authors gratefully acknowledge the financial support by the Swiss National Science Foundation (grant no. 200020_129935). We also thank the Swiss Light Source at the Paul Scherrer Institute (Villigen, Switzerland) for beamtime and their support.

Keywords E-selectin; sialyl Lewis^x; Glycomimetics; Ligand Preorganization.

The family of selectins with the three members E-, P-, and L-selectin, are key players in leukocyte trafficking in physiological and pathophysiological processes.^[1] All selectins are membrane bound C-type (*i.e.* calcium dependent) lectins that mediate the initial adhesion of leukocytes to endothelial cells. The carbohydrate binding site is located in the N-terminal lectin (Lec) domain, which is followed by the epidermal growth factor-like (EGF-like) domain, a variable number of short consensus repeats (SCR), a transmembrane region, and a C-terminal cytoplasmic tail.^[2] The three selectins share a common minimal binding motif,^[3] namely the tetrasaccharide sialyl Lewis^x (**1**, sLe^x, (Neu5Ac(α 2-3)Gal(β 1-4)[Fuc(α 1-3)]GlcNAc), Figure 1), which is presented on cell-surface glycoproteins (*e.g.* P-selectin glycoprotein ligand-1^[4] and E-selectin ligand-1^[5]).

Dysregulation of E-selectin presentation on postcapillary venules leads to excessive leukocyte recruitment and extravasation that may cause numerous pathophysiological situations, *e.g.* stroke, asthma, psoriasis, or rheumatoid arthritis.^[6] Several drug discovery programs aiming at selectin inhibition have been launched in the past.^[7] In search of drug-like antagonists for selectin-mediated diseases, sLe^x (**1**) has become the lead structure. These efforts have resulted in the development of several clinical candidates, *e.g.* CY-1503,^[8] TBC-1269^[9], or GSC-150^[10], and recently GMI-1070, a pan-selectin antagonist with a glycomimetic core,^[11] which is currently in phase 2 clinical trials for the treatment of sickle cell crisis.

The structural elucidation of the E-selectin Lec and EGF-like domains in absence of ligand^[12] and the soaked E-selectin/sLe^x (**1**) structure^[13] enhanced the understanding of the interaction on a molecular level. The comparison of both structures revealed no significant changes to the binding site residues in presence or absence of the ligand. The two moieties L-fucose (L-Fuc) and D-galactose (D-Gal) on sLe^x (**1**), as well as the carboxylate group of *N*-acetyl-D-neuraminic acid (D-Neu5Ac) form direct contacts to the protein and the Ca²⁺ ion within the binding site. The D-GlcNAc moiety and most of the D-Neu5Ac are solvent exposed. By replacing these portions of sLe^x (**1**) with hydrophobic moieties, structurally simplified and more drug-like antagonists such as CGP69669 (**2**)^[14] with fifteenfold improved affinity, as determined by isothermal titration calorimetry (ITC),^[15] were developed (Figure 1). Further investigations on substituents on the cyclohexane-diol moiety led to the closely related antagonist **3**, which exhibited a further sixfold improvement in affinity over **2** by optimizing the spatial orientation of the L-Fuc and D-Gal moieties.^[16] Finally, the introduction of a benzoyl group in the 2-hydroxyl position of D-Gal in antagonist **4**, further increased affinity

fivefold compared to **3**.^[16] However, the benefits in affinity that arise from benzylation could not be rationalized since any substituent on the 2-hydroxyl group of D-Gal is expected to protrude into the solvent assuming an identical binding mode as observed in E-selectin/sLe^x (**1**) crystal structure.^[13] In this report, we provide insights on how D-Gal benzylation increases ligand affinity by aiding the pre-organization of the carboxylic acid using a combination of molecular modeling, surface plasmon resonance (SPR), ITC, and protein-ligand crystallography.

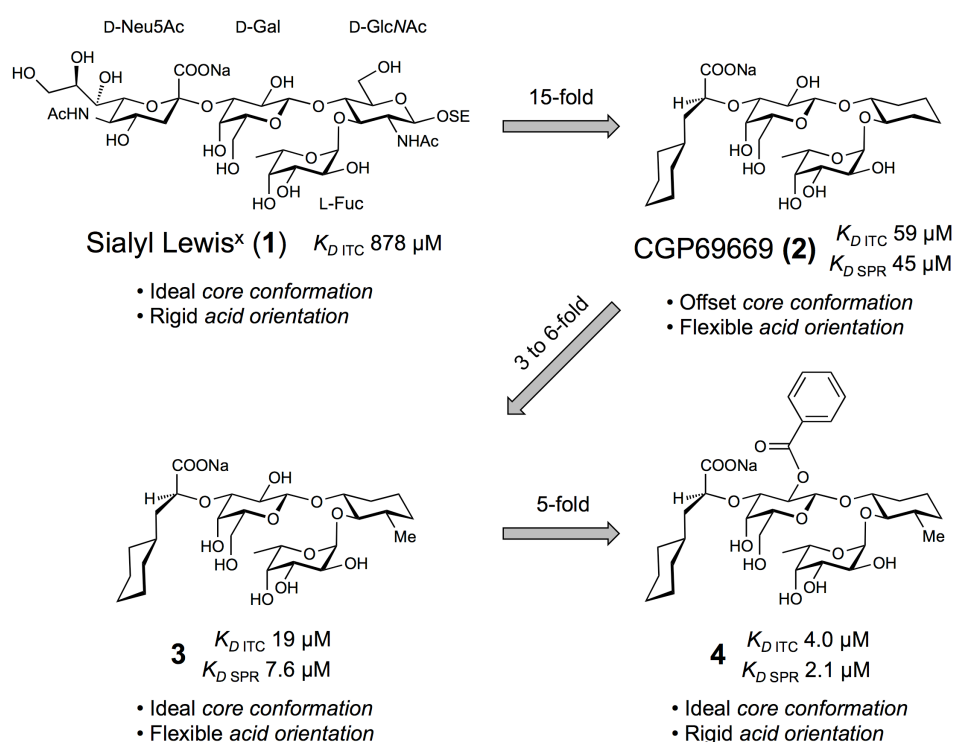


Figure 1. Series of E-selectin antagonists investigated in this study. The minimal carbohydrate binding epitope sialyl Lewis^x (sLe^x, **1**) and related glycomimetics CGP69669 (**2**), **3** and **4**. For all sLe^x mimics, D-Neu5Ac was replaced by (S)-cyclohexyl lactic acid. D-GlcNAc has been replaced by (R,R)-cyclohexane-1,2-diol (\rightarrow **2**) or (1R,2R,3S)-3-methyl cyclohexane-1,2-diol (\rightarrow **3**, **4**). Antagonist **4** is additionally benzyolated in the 2-hydroxyl position of D-Gal. K_D values from isothermal titration calorimetry ($K_{D\text{ ITC}}$) for compounds **1-3** are taken from Binder *et al.*^[15] Steady-state K_D values from surface plasmon resonance ($K_{D\text{ SPR}}$) data for compounds **2-4** are taken from Schwizer *et al.*^[16]

Two dihedral angles have previously been identified to describe the spatial organization of the pharmacophores of sLe^x (**1**). The *core conformation* describes the spatial arrangement of L-Fuc and D-Gal, while the *acid orientation* describes the orientation of the carboxylic acid group relative to the Le^x core (Figure 2).^[14b] These two torsion angles have previously been determined for sLe^x (**1**) in its protein-bound state by transferred nuclear Overhauser-NMR

(trNOE-NMR).^[17] In its bioactive conformation, the *core conformation* of sLe^x (**1**) ranges between -30° and -55°, and the *acid conformation* between 105° and 130°. ^[18] These angles were confirmed in the E-selectin structure soaked with sLe^x (**1**).^[13] Using molecular dynamic (MD) simulations, we calculated the two torsion angles in a solvated environment for sLe^x (**1**), and the three E-selectin antagonists (**2-4**). As shown in Figure 2, both, the *core conformation* and the *acid orientation* are well defined for sLe^x (**1**), since the angles are highly populated within the experimental window. This is thought to be a consequence of the exo-anomeric effect,^[19] and additionally of intramolecular hydrogen bonds and steric effects which limit the movements of the entire tetrasaccharide. The replacements of D-GlcNAc and D-NeuNAc by smaller, hydrophobic and unsubstituted residues in compound **2**, while greatly improving affinity, lead to a small offset in the *core conformation* as described by Schwizer *et al.*,^[16] and additionally introduced flexibility of carboxylic acid. While the introduction of the methyl group in the 3-position of cyclohexane-diol in compound **3** counteracted the deficit in *core conformation*, thus leading to improved affinity, the flexibility of the carboxylic acid remained similar. The introduction of the bulky benzoate in 2-position of D-Gal in compound **4** restored the pre-organization of the acid orientation to the position as determined for sLe^x (**1**).

With ligand preorganization in the bioactive conformation, two benefits are apparent: i) The association rate of binding (k_{on}) is increased since the ligand does not need to reorganize prior to binding. Indeed, the association rate of the E-selectin ligand-1 (ESL-1), a glycoprotein bearing the pre-organized sLe^x (**1**) motif, has a rapid association rate of $7.4 \cdot 10^4 \text{ M}^{-1}\text{s}^{-1}$.^[20] ii) A restrained ligand usually has decreased entropic costs compared its flexible counterpart, since the flexibility is limited prior to complex formation with the protein. We have recently shown that sLe^x (**1**) binding to E-selectin is an entropy driven process with unfavorable enthalpy, an unusual feature for carbohydrate-lectin interactions.^[21]

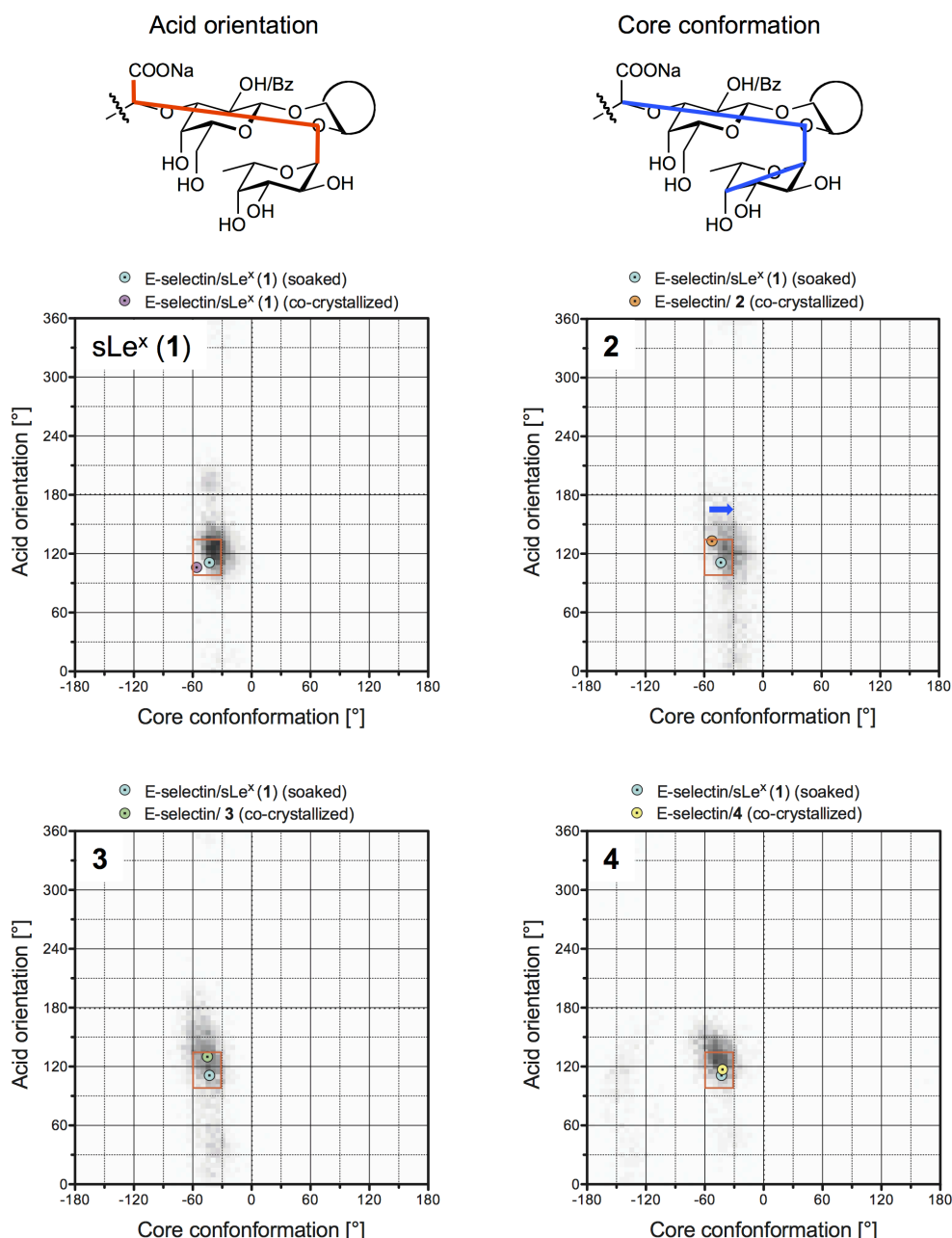


Figure 2. The *acid orientation* is a measurement of the carboxylic acid position relative to the Le^x core. The *core conformation* describes the spatial arrangement of the core itself. These torsion angles were calculated in solution for the ligands 1–4 in absence of protein with 9.6 ns MD simulations. A darker shading correlates with higher population in the respective conformation. The red rectangle encloses experimental trNOE-NMR data of sLe^x (1) bound to E-selectin.^[17] The measured angles from crystallographic structures for E-selectin soaked with sLe^x (1, cyan, PDB code 1G1T),^[13] and the co-crystallized compounds (1, magenta; 2, orange; 3, green; 4, yellow) are given. The blue arrow indicates the small offset in core-conformation of antagonist 2.

To address binding kinetics, the association and dissociation constants for antagonists 2–4 were determined with SPR (Table 1). The dissociation constants (k_{off}) for all three antagonists

were similarly rapid, which translated in short half-lives ($t_{1/2}$) of 0.5–0.8 s. The improvements in affinities therefore arise from increasing association constants. The offset in both, the *core conformation* and the *acid orientation* in compound **2** requires the ligand to adopt an energetically unfavored *core conformation* and freezing of the carboxylic. The pre-organization of the modified Le^x core in antagonist **3** lead to a tenfold faster association rate (k_{on}). Finally, the association rate was further enhanced fourfold for compound **4**, in which both, the *core conformation* and *acid orientation* are pre-organized. These observations give reason to believe that optimal orientation of the pharmacophores in sLe^x mimetics greatly enhance ligand association, while ligand unbinding remains unaltered.

Table 1: Binding kinetics of the glycomimetic E-selectin antagonists **2–4** as determined by SPR.

Ligand	K_D eq [a] [μM]	k_{on} [$10^5 \text{ M}^{-1} \text{ s}^{-1}$]	k_{off} [s^{-1}]	$t_{1/2}$ [b] [s]
2	50.4	0.2	1.0	0.7
3	6.5	2.0	1.3	0.5
4	1.5	8.5	0.9	0.8

[a] The K_D was derived from steady-state response fit to a single binding site model.

[b] The half-life ($t_{1/2}$) was calculated with $t_{1/2} = \ln 2 / k_{off}$.

To investigate whether stabilization of the *acid orientation* translates into the expected benefit in entropy, we determined the thermodynamic profile of antagonist **4** by ITC (Table 2). As previously reported, the Le^x core stabilization (**2**→**3**) lead an improvement in both, enthalpy ($\Delta\Delta H = -0.4 \text{ kJ mol}^{-1}$) and, more significantly, in entropy ($-T\Delta\Delta S = -2.4 \text{ kJ mol}^{-1}$). The introduction of the benzoate moiety (**3**→**4**) unexpectedly enhanced binding by improving enthalpy ($\Delta\Delta H = -6.6 \text{ kJ mol}^{-1}$), while showing an unexpected compensation in entropy ($-T\Delta\Delta S = +2.9 \text{ kJ mol}^{-1}$).

Table 2: Thermodynamic profile of sLe^x (**1**) and the glycomimetic antagonists **2–4** as determined by ITC.^[a]

Ligand	K_D [μM]	ΔG [kJ mol^{-1}]	ΔH [kJ mol^{-1}]	$-T\Delta S$ [kJ mol^{-1}]	N	ΔC_p [$\text{kJ mol}^{-1} \text{ K}^{-1}$]
1 ^[15]	878 ± 93	-17.5 ± 0.2	$+5.4 \pm 0.7$	-23.0 ± 1.0	1	-0.06
2 ^[15]	59 ± 4	-24.2 ± 0.2	-5.3 ± 0.4	-18.9 ± 0.6	0.93 ± 0.08	-0.11
3 ^[15]	19 ± 2	-27.1 ± 0.2	-5.8 ± 0.1	-21.3 ± 0.4	0.97 ± 0.01	n.d.
4	4.0 ± 0.3	-30.8 ± 0.2	-12.4 ± 0.6	-18.4 ± 0.8	1.00 ± 0.07	-0.16

[a] ITC data of compounds **1–3**, except for the ΔC_p values, are taken from Binder *et al.*^[15] K_D , ΔH , and the stoichiometry N for antagonist **4** were determined by ITC, while ΔG and $T\Delta S$ were calculated using following equation: $\Delta G = \Delta H - T\Delta S = RT \ln K_D$. Changes in heat capacity ΔC_p for compounds **1**, **2**, and **4** were determined from measurements at three temperatures (283.15, 298.15K, and 310.15 K). ΔC_p was calculated with following equation: $\Delta C_p = \Delta H_{T2} - \Delta H_{T1} (T2 - T1)^{-1}$.

Since the overall entropy term, as determined by ITC, is the sum of not only conformational entropy, but also solvation entropy of both the protein and the ligand ($\Delta S_{tot} = \Delta S_{confL} + \Delta S_{confP} + \Delta S_{solvL} + \Delta S_{solvP}$), it is conceivable that the introduction of a hydrophobic aromatic group in antagonist **4** might alter the solvation properties of the ligand. However, the benzoate moiety in the 2-position of D-Gal is expected to be solvent exposed after binding and is therefore not likely to undergo desolvation upon binding. Changes in heat capacity, ΔC_p , are commonly referred to as parameter for solvent reorganization, especially of hydrophobic interactions.^[22] A negative ΔC_p value correlates with a burial of non-polar surfaces. To assess changes in heat capacity for E-selectin antagonists, we determined the thermodynamic parameters of compounds **1**, **2** and **4** at three different temperatures (283.15, 298.15K, and 310.15 K). As commonly observed for carbohydrate-lectin interactions,^[21a, 23] the changes in heat capacity for sLe^x (**1**) of $-0.06 \text{ kJ mol}^{-1} \text{ K}^{-1}$ were small and negative. This observation was also consistent after introducing hydrophobic portions in antagonists **2** ($-0.11 \text{ kJ mol}^{-1} \text{ K}^{-1}$) and **4** ($-0.16 \text{ kJ mol}^{-1} \text{ K}^{-1}$). These slight differences were within the experimental range. The small increase in ΔC_p indicate that there was a minimally increase in burial of hydrophobic surfaces from sLe^x (**1**) to antagonists **2** and **4**.

For further insights into E-selectin-ligand interactions, especially of antagonist **4** in relation to antagonist **3**, we co-crystallized compounds **1–4** with E-selectin consisting of the four N-terminal domains (lectin domain, EGF-like domain, and first two short consensus repeats). Compared to the previously reported crystallographic structures of E-selectin without ligand or soaked with sLe^x (**1**), we have recently demonstrated with sLe^x (**1**) and antagonist **3**, that E-selectin-ligand co-crystallization reveals an alternative overall conformation of the protein.^[Manuscript 1] We further showed that small-molecule–ligand interactions lead to an induced fit of the protein in solution and an overall extension of the protein, which involves rearrangements over the entire length of the lectin domain to the EGF-like domain interface (Figure 2). This extension is described by a 30° increase in the angle from the Ca²⁺ ion to Trp1_{Cα} at the pivot to Cys144_{Cα} within the EGF-like domain (Figure 2a). Three major changes in the direct vicinity of the ligand were identified: i) The loop consisting of the amino acid residues 81–89 was entirely rearranged and lead to a constriction around the Ca²⁺ ion an the L-Fuc moiety. Additional direct hydrogen bonds between the 2-hydroxyl group of L-Fuc and Gln85 and Glu88 were formed. ii) Rearrangement also allowed for Glu107 to form a hydrogen bond with the 3-hydroxyl group of L-Fuc. iii) Co-crystallization of antagonist **3** showed that Pro46 and its neighboring residues shift towards the D-Gal moiety, allowing for

a water-mediated interaction to the 6-hydroxyl group of D-Gal. Overall, these three modifications upon an induced fit allowed for a more “pocket-like” binding site topology compared to the soaked E-selectin/sLe^x (1) structure.^[12] While the protein underwent significant alterations when co-crystallized, the binding mode and the general orientation of pharmacophores of sLe^x (1) and antagonist 3 remained similar to the soaked structure. Here we report crystallographic structures of antagonists 2 and 4. Antagonist 2 crystallized in similar conditions as sLe^x (1) as the mimic 3, and showed similar overall conformation of the protein. For antagonist 4 however, different crystallization conditions were required to yield diffraction quality crystals, which resulted in alternative crystal packing. Nonetheless, the 30° increase in the Ca²⁺–Trp1_{Cα}–Cys144_{Cα} angle as observed in co-crystallized structures with ligands 1–3 was conserved. Furthermore, the specific alterations regarding the protein backbone surrounding the binding site were also in agreement (Figure 2b). The unexpected loss in entropy upon pre-organization (3→4) can therefore not be attributed to a different binding mode of antagonist 4 or major differences in the conformation of E-selectin.

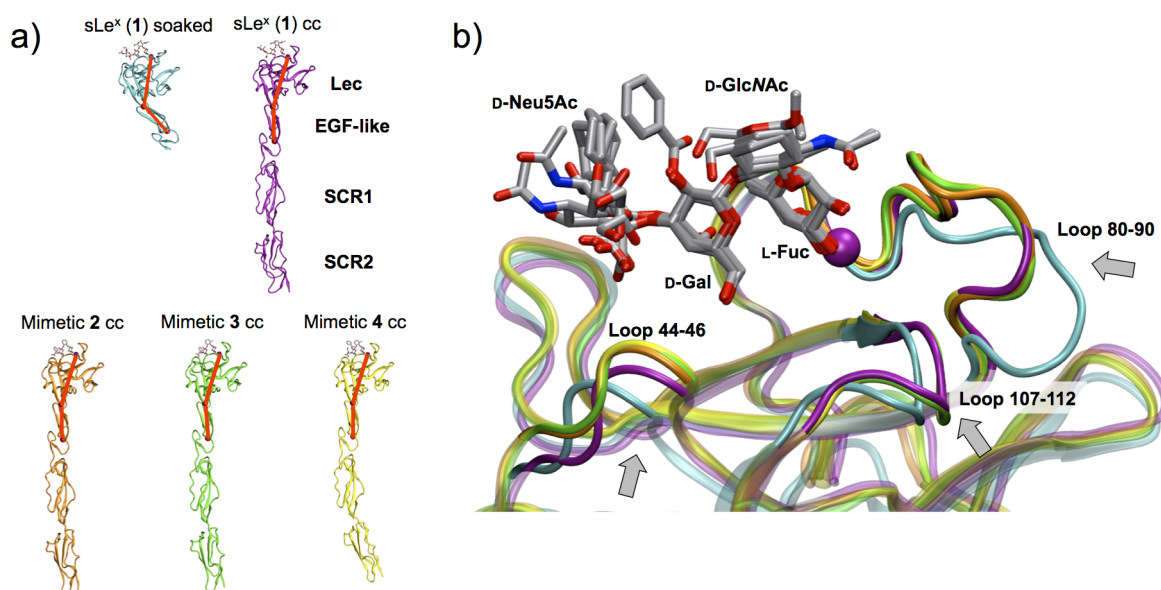


Figure 2. Structures of E-selectin soaked with sLe^x (1, cyan, PDB code 1G1T)^[13] or co-crystallized (cc) with sLe^x (1) (magenta), antagonists 2 (orange), 3 (green), and 4 (yellow). The Ca²⁺ ion is depicted as purple sphere. a) Overall conformation of the E-selectin structures. The Ca²⁺–Trp1_{Cα}–Cys144_{Cα} angle is marked red and was consistently increased by ~30° for all co-crystallized structures compared to the soaked structure. b) Superposition of the protein backbone surrounding the ligand. Regions involved in major alterations from the soaked to the co-crystallized structure are marked by an arrow. The Le^x core of sLe^x (1) is well aligned between the soaked and co-crystallized structure, as well as for antagonists 2–4. Flexibility of the D-Neu5Ac/(S)-cyclohexyl lactic acid is apparent.

Detailed inspection of the (*S*)-cyclohexyl lactic acid moiety and the side-chain rotamers of amino acid residues within the binding site, however, revealed significant alterations. While most binding site residues, especially those interacting with the L-Fuc and D-Gal moieties, have identical orientations (Figure 3a), significant changes are observed for the two residues Glu98 and Lys99 (Figure 3b). The benzoate on D-Gal in antagonist **4** provides for sigma-pi interactions to the cyclohexane moiety with a distance of 3.7 Å of the closest heavy atom to the centre of the aromatic ring. This intramolecular interaction shifts the cyclohexane moiety into a slightly different position, as indicated by the arrow in Figure 3b. The absence of the aromatic benzoyl group allows for more flexibility of the (*S*)-cyclohexyl lactic acid in antagonists **2** and **3**. Indeed, comparing the crystallographic b-factors, which give an indication for thermal motion of the cyclohexyl moiety for antagonists **3** and **4** to their respective modified Le^x core, the values are significantly higher in antagonist **3**, while being similar to the modified Le^x core in antagonist **4**. This is also true for b-values for the carboxylic acid, which are also increased in compound **3**. Thermodynamically speaking, a more steady orientation of the carboxylic acid might allow for higher quality and energetically improved interactions of the carboxylic acid to Tyr48 and Arg97 and might therefore account for some of the increase in enthalpy. The interplay of the cyclohexane ring and the benzoate in compound **4** prior to binding that leads to rigidification of the molecule is expected lead to beneficial entropy. Concerning the binding site, differences in the side-chain orientation of Glu98 and Lys99 were observed, while being conserved along the backbone. The side-chain atoms of these two residues in the structure of antagonist **3** exhibit greater b-values compared to their backbone atoms, suggesting flexibility. With compound **4**, the b-values of Glu98/Lys99 side-chain atoms are, again in comparison to their respective backbone atoms, not increased to the extent observed with antagonist **3**. Furthermore, different side-chain rotamers are adopted that allowed the cyclohexane ring to be accommodated between the two residues. This provided for hydrophobic interactions, as indicated by distances of the cyclohexane ring of 3.7 Å and 3.9 Å to Glu98_{C γ} and Lys99_{C γ} , respectively. These interactions might benefit from the slight reorientation and stabilization of the cyclohexane moiety, which is likely caused by the intramolecular sigma-pi interaction. By interacting with the cyclohexane moiety in antagonist **4**, their flexibility is limited which might account for loss in entropy. Furthermore, such hydrophobic interactions require highly ordered water molecules surrounding the surfaces to be released into bulk water upon binding, a process that is accompanied by a large gain in entropy.^[24] These classical hydrophobic interactions are accompanied by small favorable enthalpic contributions that are

strongly temperature dependent (ΔC_p). In contrast, non-classical hydrophobic interactions are largely driven by enthalpic contributions.^[22b] Here, water molecules surrounding hydrophobic surfaces form strong interactions to bulk water upon desolvation,^[22c] especially for water molecules in hydrophobic cavities, where they are not well ordered. Such a non-classical hydrophobic behavior has previously been observed for ligands of increasing lipophilicity.^[22c, 25] However, no data on surface-exposed binding sites exist.

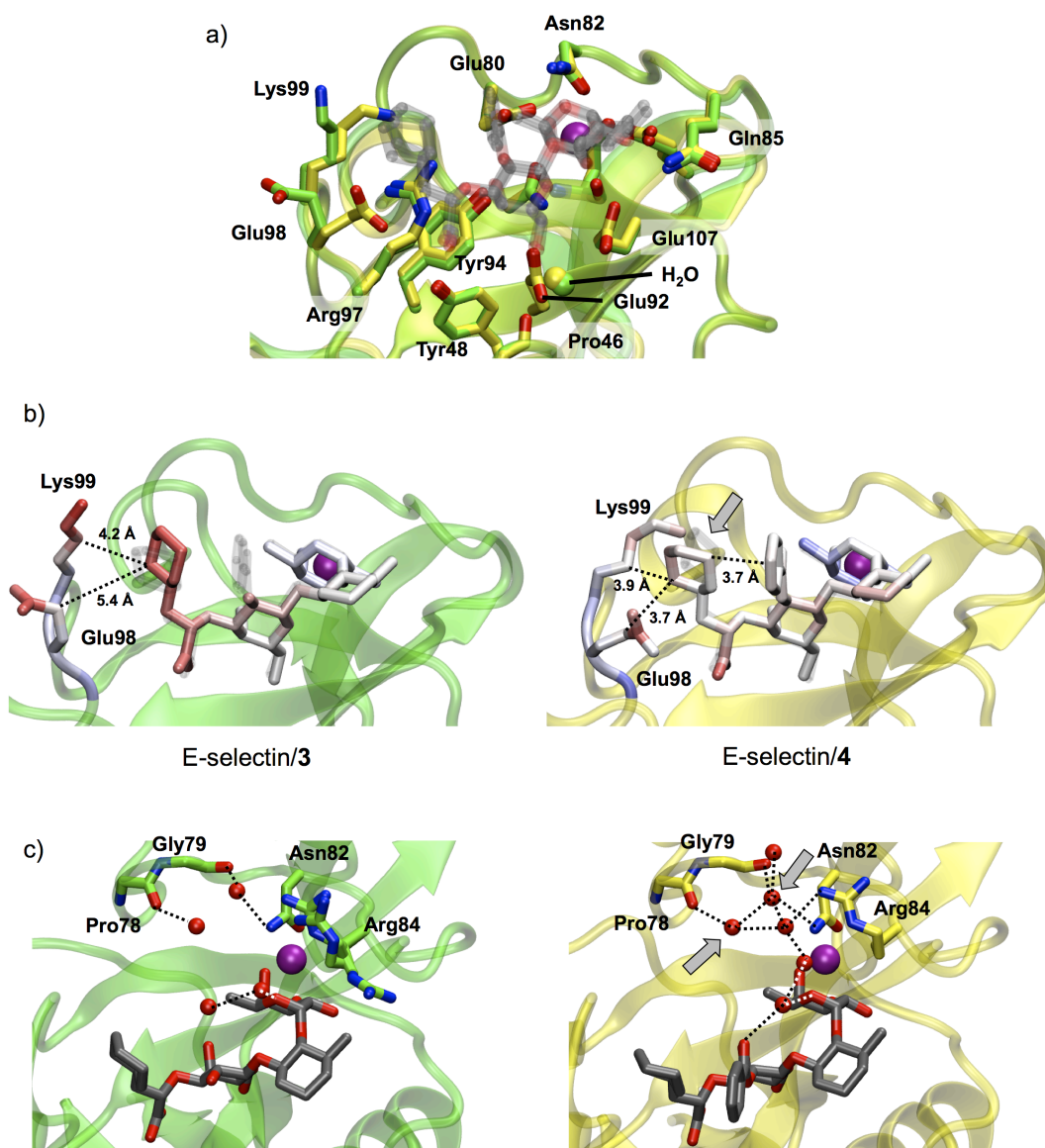


Figure 3. E-selectin co-crystallized with the flexible antagonists **3** (green), and the pre-organized antagonist **4** (yellow). a) Comparison of binding site residues in direct or water-mediated (Pro46) contact to the ligands (transparent). All residues are in identical orientations, except for Glu89 and Lys99 side-chains. The water mediating an interaction to Pro46 is shown as green or yellow sphere. The Ca^{2+} ion is depicted as purple sphere. b) Comparison of crystallographic b-values in blue (low b-value), white (average), and red (high). For the Glu98 and Lys99 side-chains, the b-values were normalized against their backbone atoms. For each ligand, the b-values of the (S)-cyclohexyl lactic

acid were normalized against the modified Le^x core. The arrow indicates the shift of the cyclohexane ring orientation. c) Water clustering surrounding the benzoate moiety of antagonist **4** and the Arg84 side-chain. Connected lines are drawn for distances up to 3.2 Å. Water molecules with near-identical positions in the structures with antagonist **3** or **4** are marked by arrows.

Finally, the Arg84 residue, belonging to the 81–89 loop involved in the induced fit, is shifted towards ligand without directly contributing to binding.^[Manuscript 1] In the co-crystal structure with antagonist **3**, the Arg84 side-chain adopts various rotamers, thus suggesting flexibility. With antagonist **4**, however, the Arg84 is in a well defined orientation, and is surrounded by a cluster of water molecules which extend to the carboxylic oxygen of the benzoate moiety (Figure 3c). This water cluster might lead to a stabilization of the guanidinium moiety of Arg84, thus further being detrimental for entropy. However, the importance of such effects that propagate over several water molecules are difficult to quantify.

The above observations regarding the flexibility of (*S*)-cyclohexyl lactic acid, Arg84, Glu98, and Lys99 were also observed in the structure with antagonist **2** (not shown), further suggesting that these effects are indeed a direct consequence of D-Gal benzylation.

In summary, we provide kinetic, thermodynamic, and crystallographic insights for a series of glycomimetic E-selectin antagonists. sLe^x (**1**) is a rigid carbohydrate with optimal *core conformation* and rigid *acid orientation*. Substitution of D-GlcNAc and D-Neu5Ac with (1*R*,2*R*)-cyclohexane-1,2-diol and (*S*)-cyclohexyl lactic acid (\rightarrow **2**), respectively, results in improved affinity due to alteration of solvation properties.^[15] However, this antagonist suffers from an offset in *core conformation* and from flexibility of the *acid orientation*.^[16] Methylation of the 3-position of cyclohexane-diol (\rightarrow **3**) corrects the deficit in *core conformation*, while the *acid orientation* remains flexible. Finally, we propose that benzylation of the 2-position of D-Gal (\rightarrow **4**), pre-organizes the *acid orientation* by forming sigma-pi interactions with the cyclohexane ring. While ligand pre-organization lead to the expected increase in association kinetics, comparison of the thermodynamic profiles of antagonist **3** \rightarrow **4** revealed a significant increase in enthalpy accompanied by an unexpected loss in entropy. Such paradox entropic behavior has previously been reported for the flexible and constrained ligands to the SH2 domain of the growth factor receptor-bound protein 2 (Grb2),^[26] where more interactions and closer contact of the protein with the pre-organized ligands were observed. Based on co-crystallization of E-selectin antagonists **2–4**, we observed that: i) intramolecular interactions in antagonist **4** lead to a more rigid orientation of the *acid orientation*, ii) the two residues Glu98 and Lys99 are rigidified and form

hydrophobic interactions, and iii) water clustering occurs surrounding the benzoate group. While quantitative dissection of each of these observations is troublesome, we provide a qualitative explanation for the entropic paradox of ligand pre-organization of antagonist **4**. The fact that E-selectin undergoes an induced fit upon binding also needs to be taken into consideration. Higher ligand affinity could lead to a more stable adoption of the induced fit conformation, which might also be reflected in the thermodynamic profile. Also such effects are difficult to quantify and are largely elusive.

Regarding kinetics of glycomimetic small-molecule E-selectin antagonists, the major future challenge is to improve half-lives ($t_{1/2}$), since this often can be directly translated into compound efficacy.^[27] With structural information on the induced-fit conformation of the lectin domain, rational drug design can now move forward towards high-affinity antagonists with long half-lives.

References

- [1] H. Lis, N. Sharon, *Chem. Rev.* **1998**, *98*, 637-674.
- [2] D. V. Erbe, B. A. Wolitzky, L. G. Presta, C. R. Norton, R. J. Ramos, D. K. Burns, J. M. Rumberger, B. N. N. Rao, C. Foxall, B. K. Brandley, L. A. Lasky, *J. Cell Biol.* **1992**, *119*, 215-227.
- [3] a) M. L. Phillips, E. Nudelman, F. C. Gaeta, M. Perez, A. K. Singhal, S. Hakomori, J. C. Paulson, *Science* **1990**, *250*, 1130-1132; b) M. J. Polley, M. L. Phillips, E. Wayner, E. Nudelman, A. K. Singhal, S. Hakomori, J. C. Paulson, *Proc. Natl. Acad. Sci. U. S. A.* **1991**, *88*, 6224-6228; c) G. Walz, A. Aruffo, W. Kolanus, M. Bevilacqua, B. Seed, *Science* **1990**, *250*, 1132-1135.
- [4] a) K. L. Moore, N. L. Stults, S. Diaz, D. F. Smith, R. D. Cummings, A. Varki, R. P. McEver, *J. Cell Biol.* **1992**, *118*, 445-456; b) D. Sako, X. J. Chang, K. M. Barone, G. Vachino, H. M. White, G. Shaw, G. M. Veldman, K. M. Bean, T. J. Ahern, B. Furie, D. A. Cumming, G. R. Larsen, *Cell* **1993**, *75*, 1179-1186.
- [5] a) M. Lenter, A. Levinovitz, S. Isenmann, D. Vestweber, *J. Cell Biol.* **1994**, *125*, 471-481; b) A. Levinovitz, J. Muhlhoff, S. Isenmann, D. Vestweber, *J. Cell Biol.* **1993**, *121*, 449-459.
- [6] S. A. Mousa, D. A. Cheresch, *Drug Discov. Today* **1997**, *2*, 187-199.
- [7] a) B. Ernst, J. L. Magnani, *Nat. Rev. Drug Discov.* **2009**, *8*, 661-677; b) N. Kaila, B. E. Thomas, *Med. Res. Rev.* **2002**, *22*, 566-601.
- [8] K. M. Kerr, W. R. Auger, J. J. Marsh, R. M. Comito, R. L. Fedullo, G. J. Smits, D. P. Kapelanski, P. F. Fedullo, R. N. Channick, S. W. Jamieson, K. M. Moser, *Am. J. Respir. Crit. Care Med.* **2000**, *162*, 14-20.
- [9] T. P. Kogan, B. Dupre, H. Bui, K. L. McAbee, J. M. Kassir, I. L. Scott, X. Hu, P. Vanderslice, P. J. Beck, R. A. F. Dixon, *J. Med. Chem.* **1998**, *41*, 1099-1111.
- [10] Y. Wada, T. Saito, N. Matsuda, H. Ohmoto, K. Yoshino, M. Ohashi, H. Kondo, H. Ishida, M. Kiso, A. Hasegawa, *J. Med. Chem.* **1996**, *39*, 2055-2059.
- [11] J. S. Chang, J. T. Patton, A. Sarkar, B. Ernst, J. L. Magnani, P. S. Frenette, *Blood* **2010**, *116*, 1779-1786.
- [12] B. J. Graves, R. L. Crowther, C. Chandran, J. M. Rumberger, S. Li, K. S. Huang, D. H. Presky, P. C. Familletti, B. A. Wolitzky, D. K. Burns, *Nature* **1994**, *367*, 532-538.
- [13] W. S. Somers, J. Tang, G. D. Shaw, R. T. Camphausen, *Cell* **2000**, *103*, 467-479.
- [14] a) K. E. Norman, G. P. Anderson, H. C. Kolb, K. Ley, B. Ernst, *Blood* **1998**, *91*, 475-483; b) H. C. Kolb, B. Ernst, *Chem. Eur. J.* **1997**, *3*, 1571-1578; c) H. C. Kolb, B. Ernst, *Pure Appl. Chem.* **1997**, *69*, 1879-1884.
- [15] F. P. Binder, K. Lemme, R. C. Preston, B. Ernst, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 7327-7331.
- [16] D. Schwizer, J. T. Patton, B. Cutting, M. Smiesko, B. Wagner, A. Kato, C. Weckerle, F. P. Binder, S. Rabbani, O. Schwardt, J. L. Magnani, B. Ernst, *Chemistry* **2012**, *18*, 1342-1351.
- [17] a) R. M. Cooke, R. S. Hale, S. G. Lister, G. Shah, M. P. Weir, *Biochemistry (Mosc.)* **1994**, *33*, 10591-10596; b) R. Harris, G. R. Kiddle, R. A. Field, M. J. Milton, B. Ernst, J. L. Magnani, S. W. Homans, *J. Am. Chem. Soc.* **1999**, *121*, 2546-2551; c) P. Hensley, P. J. Mcdevitt, I. Brooks, J. J. Trill, J. A. Feild, D. E. McNulty, J. R. Connor, D. E. Griswold, N. V. Kumar, K. D. Kopple, S. A. Carr, B. J. Dalton, K. Johanson, *J. Biol. Chem.* **1994**, *269*, 23949-23958.
- [18] a) M. Rinnbauer, B. Ernst, B. Wagner, J. Magnani, A. J. Benie, T. Peters, *Glycobiology* **2003**, *13*, 435-443; b) K. Scheffler, J. R. Brisson, R. Weisemann, J. L. Magnani, W. T. Wong, B. Ernst, T. Peters, *J. Biomol. NMR* **1997**, *9*, 423-436; c) K.

- Scheffler, B. Ernst, A. Katopodis, J. L. Magnani, W. T. Wang, R. Weisemann, T. Peters, *Angew. Chem. Int. Ed. Engl.* **1995**, *34*, 1841-1844.
- [19] I. Tvaroska, T. Kozar, *J. Am. Chem. Soc.* **1980**, *102*, 6929-6936.
- [20] M. K. Wild, M. C. Huang, U. Schulze-Horsel, P. A. van der Merwe, D. Vestweber, *J. Biol. Chem.* **2001**, *276*, 31602-31612.
- [21] a) M. Ambrosi, N. R. Cameron, B. G. Davis, *Org. Biomol. Chem.* **2005**, *3*, 1593-1608; b) T. K. Dam, C. F. Brewer, *Chem. Rev.* **2002**, *102*, 387-429.
- [22] a) N. T. Southall, K. A. Dill, A. D. J. Haymet, *J. Phys. Chem. B* **2002**, *106*, 521-533; b) S. W. Homans, *Drug Discov. Today* **2007**, *12*, 534-539; c) E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem. Int. Ed.* **2003**, *42*, 1210-1250.
- [23] E. J. Toone, *Curr. Opin. Struct. Biol.* **1994**, *4*, 719-728.
- [24] C. Tanford, *Science* **1978**, *200*, 1012-1018.
- [25] a) L. Englert, A. Biela, M. Zayed, A. Heine, D. Hangauer, G. Klebe, *Bba-Gen Subjects* **2010**, *1800*, 1192-1202; b) T. Young, R. Abel, B. Kim, B. J. Berne, R. A. Friesner, *Proc. Natl. Acad. Sci. U. S. A.* **2007**, *104*, 808-813.
- [26] a) S. F. Martin, *Pure Appl. Chem.* **2007**, *79*, 193-200; b) A. P. Benfield, M. G. Teresk, H. R. Plake, J. E. DeLorbe, L. E. Millspaugh, S. F. Martin, *Angew. Chem. Int. Ed.* **2006**, *45*, 6830-6835; c) J. H. Clements, J. E. DeLorbe, A. P. Benfield, S. F. Martin, *Acta Crystallogr D* **2010**, *66*, 1101-1115; d) J. E. DeLorbe, J. H. Clements, M. G. Teresk, A. P. Benfield, H. R. Plake, L. E. Millspaugh, S. F. Martin, *J. Am. Chem. Soc.* **2009**, *131*, 16758-16770.
- [27] R. A. Copeland, D. L. Pompliano, T. D. Meek, *Nat. Rev. Drug Discov.* **2006**, *5*, 730-739.

Supporting Information

Acid Pre-Organization of a Glycomimetic E-selectin Antagonist

Roland C. Preston, Katrin Lemme, Céline Weckerle, Martin Smieško, Roman Jakob, Timm Maier, Beat Ernst

*Institute of Molecular Pharmacy, Pharmacenter, University of Basel, Klingelbergstrasse 50,
CH-4056 Basel, Switzerland.*

and

Biocenter, University of Basel, Klingelbergstrasse 70, 4056 Basel, Switzerland.

- 1. MD simulations of solution conformations**
- 2. Synthesis of antagonists 2-4**
- 3. Protein expression and purification**
- 4. Surface plasmon resonance (SPR)**
- 5. Isothermal titration calorimetry (ITC)**
- 6. E-selectin co-crystallization with compounds 2 and 4**
- 7. Structure determination and refining**

1. MD simulations of solution conformations

The mixed torsional/low-mode sampling algorithm implemented in MacroModel^[S1] (version 9.8107) in conjunction with the OPLS 2005 force-field^[S2] was used to find the global minima of ligands in implicit solvent conditions (GB/SA, water). The global minimum conformation was then positioned in the center of an orthorhombic periodic boundary system extending by 12.0 Å beyond the ligand structure in each (*x*, *y*, *z*) direction and surrounded by the TIP3P water molecules. Sodium counterions were added in order to neutralize total charge of the system to zero. The system was then minimized using Desmond simulation software (provenance 2.3.4.1).^[S3] Simulated annealing (SA) at constant volume was employed in order to provide the compounds an option to readjust their conformation as well as to organize surrounding solvent molecules for optimal interaction. SA setup details (Desmond): for the initial period of 30 ps the system was equilibrated at 0 K. From 30 to 300 ps the temperature was gradually increased to 400 K, from 300 to 500 ps maintained at 400 K and from 500 to 1000 ps the system was slowly cooled down to 300 K followed by the final equilibration from 1000 to 1200 ps. After this preparatory step a 9.6 ns long production phase molecular dynamics (MD) simulation was then performed in Desmond. Snapshots of the system were taken every 2.4 ps, thus overall 4000 frames were saved from the production phase. For each saved frame the two most important structural descriptors – the *core conformation* and the *acid orientation* of the ligand – were evaluated. The distribution of the to descriptor values was plotted in a gradient-colored histogram.

2. Synthesis of antagonists 2-4

Antagonist {(1*R*,2*R*)-2-[(α -L-Fucopyranosyl)oxy]-cyclohex-1-yl} 3-*O*-[sodium (1*S*)-1-carboxy-2-cyclohexyl-ethyl]- β -D-galactopyranoside (**2**) was synthesized according to Kolb and Ernst.^[S4] Antagonists {(1*R*,2*R*,3*S*)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl} 3-*O*-[sodium (1*S*)-1-carboxy-2-cyclohexyl-ethyl]- β -D-galactopyranoside (**3**) and {(1*R*,2*R*,3*S*)-2-[(α -L-Fucopyranosyl)oxy]-3-methyl-cyclohex-1-yl} 2-*O*-benzoyl-3-*O*-[sodium (1*S*)-1-carboxy-2-cyclohexyl-ethyl]- β -D-galactopyranoside (**4**) were synthesized described by Schwizer *et al.*^[S5]

3. Protein expression and purification

Human E-selectin consisting of the lectin domain, the EGF-like domain and the six short consensus repeats fused to the Fc-part of human IgG1 (E-selectin_{LEC6}-IgG_{Fc}) was expressed and purified according to Binder *et al.*^[S6] This construct was used for ITC and SPR experiments. For co-crystallization, a deglycosylated monomeric construct consisting of the lectin domain, EGF-like domain and the first two short consensus repeats (E-selectin_{LEC2}) was expressed and purified according to Preston *et al.*^[Manuscript 1]

4. Surface plasmon resonance (SPR)

SPR measurements of antagonists **2–4** were performed as previously described.^[S5] Briefly, a research grade CM5 sensor chip (GE Healthcare, Freiburg, Germany) was functionalized with polyclonal anti-human IgG_{Fc} antibody (Sigma Aldrich GmbH, Buchs, Switzerland) via amine coupling according to the protocol of the manufacturer and subsequently with E-selectin_{LEC6}-IgG_{Fc}. A reference cell was not functionalized with E-selectin. Ligands were injected in HBS-P buffer (10 mM HEPES, pH 7.4, 150 mM NaCl, 0.002% v/v surfactant P20) supplemented with 20 mM CaCl₂ and 2.5 % DMSO. Steady-state K_D , association (k_{on}) and dissociation constants (k_{off}) were determined with the Scrubber 2.0a software (BioLogic Software). All experiments were run on a Biacore 3000 instrument (Biacore, Uppsala, Sweden). Half-life values ($t_{1/2}$) were calculated with equation S1.

$$t_{1/2} = \ln(2/k_{off}) \quad \text{Eq. S1}$$

5. Isothermal titration calorimetry (ITC)

ITC measurements of compound **4** were performed and analyzed at 298.15 K as previously reported.^[S6] The measurement was performed in triplicate with protein (E-selectin_{LEC6}-IgG_{Fc}) concentrations between 30–45 μ M, and ligand concentrations of 475–1000 μ M. For measurements of heat capacities (ΔC_p) of compounds **1**, **2**, and **4**, measurements were performed at 283.15 K and 310.15 K. Together with data from measurements at 298.15 K, the enthalpy (ΔH) was calculated using equation S2.

$$\Delta C_p = \frac{\Delta H^\circ_{T2} - \Delta H^\circ_{T1}}{(T_2 - T_1)} \quad \text{Eq. S2}$$

6. E-selectin co-crystallization with compounds **2** and **4**

E-selectin_{LEC2} was expressed in kifunensine-treated Chinese hamster ovary (CHO) cells, which yielded active, high-mannosylated protein as previously described.^[Manuscript 1] The high-mannose *N*-linked glycans were cleaved with endoglycosidase H_f and concentrated in 10 mM Tris-HCl, 5 mM CaCl₂ to approximately 50 mg/mL. E-selectin-ligand co-crystals were grown at 20 °C in sitting-drop vapor diffusion. Initial crystals for E-selectin/**2** were obtained at 25 mg/ml protein with tenfold molar excess (10 mM) of antagonist **2** in 23 % v/v PEG8000, 0.1 M MOPS pH 6.2, and 0.2 M CaCl₂ within 24 hours. Final crystals for diffraction measurements were grown in decreased precipitation concentrations (13.5 % PEG8000) under the same conditions after microseeding. For antagonist **4** co-crystallization, E-selectin_{LEC2} (25 mg/ml) was incubated with a 10 mM precipitated solution of the ligand at 37 °C for 2 h. Remaining ligand precipitation was removed by brief centrifugation. With the same conditions as for antagonist **3**, antagonist **4** yielded only thin and bent crystals. Initial co-crystals with antagonist **4** were grown in 20 % PEG8000, 0.1 M Tris-HCl pH 8.0, and 0.2 M LiCl within 48 hours. Crystal conditions were optimized with microseeding to 15 % PEG8000, 0.1 M glycine pH 9.5, and 1.0 M LiCl. Both, co-crystals with antagonist **2** and **4** were flash cooled to 100 K with perfluoropolyether (Hampton Research, CA) as cryoprotectant. Data was collected with synchrotron radiation ($\lambda = 0.99987$ Å for antagonist **3**, $\lambda = 1.00000$ Å for antagonist **4**) at the PXI (**3**) or PXIII (**4**) beamline at the Swiss Light Source, Switzerland.

7. Structure determination and refining

Diffraction images were indexed, processed, and scaled with the XDS package.^[S7] The structures were solved by molecular replacement with PHASER^[S8] using the structure of E-selectin/**3**.^[Manuscript 1] The structures were refined with the BUSTER software^[S9] and adjusted with the COOT software.^[S10] Geometric restraints of antagonists **2** and **4** were generated with the PRODRG server.^[S11] In the E-selectin/**3** co-crystal, 10 (1.82 %) residues are Ramachandran outliers, while 28 (5.1 %) are in the allowed region. In the E-selectin/**4** co-crystal structure, 12 residues (2.3 %) are Ramachandran outliers, while 10 (1.9 %) are in the allowed region. The statistics are given in Table S1. Figures were generated with the Visual Molecular Dynamics (VMD) software.^[S12]

Table S1. Data collection and refinement statistics.

	E-selectin/2	E-selectin/4
Data collection		
Space group	P_1	P_1
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	52.3, 58.9, 59.2	51.0, 58.2, 65.7
α , β , γ (°)	75.9, 86.9, 86.5	83.3, 80.9, 90.4
Resolution (Å)	23.58– 2.05 (2.10–2.05)	29.92–1.90 (1.95–1.90)
CC (1/2)	99.9 (91.4)	99.8 (76.7)
<i>I</i> / σI	15.1 (2.1)	12.1 (2.2)
Completeness (%)	92.4 (92.4)	96.3 (91.9)
Redundancy	1.8 (1.8)	3.4 (3.2)
Refinement		
Resolution (Å)	23.58– 2.05 (2.10–2.05)	29.92–1.90 (1.95–1.90)
No. reflections	39578	56270
R_{work} / R_{free}	22.4 / 25.6	19.1 / 23.3
No. atoms		
Protein	4380	4455
Ligand/ion	469	855
Water	191	599
<i>B</i> -factors		
Protein	62.8	30.6
Ligand/ion	79.4	42.8
Water	58.4	36.6
R.m.s. deviations		
Bond lengths (Å)	0.010	0.010
Bond angles (°)	1.24	1.19

Single crystals were used for structure determination. Values in parentheses are for highest-resolution shell.

References

- [S1] Schrödinger, in *MacroModel*, 9.8 ed., LLC, New York, NY, **2010**.
- [S2] G. A. Kaminski, R. A. Friesner, J. Tirado-Rives, W. L. Jorgensen, *J. Phys. Chem. B* **2001**, *105*, 6474-6487.
- [S3] Schrödinger, in *Desmond Molecular Dynamics System*, 3.1 ed., D.E. Shaw Research, New York, NY, **2012**.
- [S4] H. C. Kolb, B. Ernst, *Chem. Eur. J.* **1997**, *3*, 1571-1578.
- [S5] D. Schwizer, J. T. Patton, B. Cutting, M. Smiesko, B. Wagner, A. Kato, C. Weckerle, F. P. Binder, S. Rabbani, O. Schwardt, J. L. Magnani, B. Ernst, *Chemistry* **2012**, *18*, 1342-1351.
- [S6] F. P. Binder, K. Lemme, R. C. Preston, B. Ernst, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 7327-7331.
- [S7] W. Kabsch, *J Appl Crystallogr* **1993**, *26*, 795-800.
- [S8] A. J. McCoy, *Acta Crystallogr D* **2007**, *63*, 32-41.
- [S9] E. Blanc, P. Roversi, C. Vornrhein, C. Flensburg, S. M. Lea, G. Bricogne, *Acta Crystallogr D* **2004**, *60*, 2210-2221.
- [S10] P. Emsley, K. Cowtan, *Acta Crystallogr D* **2004**, *60*, 2126-2132.
- [S11] A. W. Schuttelkopf, D. M. F. van Aalten, *Acta Crystallogr D* **2004**, *60*, 1355-1363.
- [S12] W. Humphrey, A. Dalke, K. Schulten, *J. Mol. Graph. Model.* **1996**, *14*, 33-38.

Chapter 2.5. – Publication 2

Implications of the E-selectin S128R Mutation on Drug Discovery

Published in *Glycobiology* vol. 24 no. 7 pp. 592-601, 2014.

Contributions

- | | |
|-----------------|---|
| R. C. Preston | <ul style="list-style-type: none">• Molecular cloning.• Generation of stably transfected CHO cells.• Protein expression and purification.• Development and performance of target- and cell-based assays.• Manuscript preparation. |
| S. Rabbani | <ul style="list-style-type: none">• Project planning and supervision. |
| F. P. C. Binder | <ul style="list-style-type: none">• Synthesis of 3'-sialyl lactosamine. |
| S. Moes | <ul style="list-style-type: none">• LC/MS/MS analysis. |
-

Implications of the E-selectin S128R mutation for drug discovery

Roland C Preston², Said Rabbani², Florian P C Binder²,
Suzette Moes³, John L Magnani⁴, and Beat Ernst^{1,2}

²Institute of Molecular Pharmacy, Pharmacenter; ³Department of Biochemistry, Biocenter, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland; and ⁴GlycoMimetics, Inc., Gaithersburg, MD 20879, USA

Received on February 28, 2014; revised on February 28, 2014; accepted on March 26, 2014

The C-type lectin E-selectin mediates the rolling of circulating leukocytes on vascular endothelial cells during the inflammatory process. In numerous studies, the S128R mutation of the E-selectin was associated with cardiovascular and autoimmune diseases. There is evidence that the S128R E-selectin mutation leads to a loss in ligand specificity, thus increasing leukocyte recruitment. Apart from the natural tetrasaccharide ligand sialyl Lewis^x (sLe^x), it has previously been proposed that non-fucosylated carbohydrates also bind to S128R E-selectin. To evaluate the therapeutic potential of the antagonism of the E-selectin mutant, ligand specificity was reinvestigated on a molecular basis. We determined the ligand specificity of wild-type and S128R E-selectin in a target-based competitive assay, a glycan array screen and cell-based binding assays under static and flow conditions. Regarding ligand-specificity, the binding properties of S128R E-selectin were identical to those of wt E-selectin, i.e., no mutant-specific binding of 3'-sialyl-N-acetyllactosamine, heparin, fetuin and K562 cells was observed. Additionally, the binding affinities of glycomimetic E-selectin antagonists were identical for wt and S128R E-selectin. Overall, the previous reports on carbohydrate ligand promiscuity of S128R E-selectin could not be confirmed.

Keywords: E-selectin / selectin antagonists / single-nucleotide polymorphism / sLe^x mimic / S128R

Introduction

In case of inflammation, the human body possesses a highly sophisticated defense line. At the initial stage, the members of the selectin family, E-, P- and L-selectin, play an important role by recruiting leukocytes to the inflammatory site (Kansas

1996). However, excessive extravasation of leukocytes leads to tissue damage being deleterious for numerous diseases with an inflammatory component (Ernst and Magnani 2009). Blocking selectins with selectin antagonists is, therefore, regarded as a promising anti-inflammatory treatment.

Selectins are C-type (calcium dependent) lectins consisting of a C-terminal cytoplasmic domain followed by a transmembrane region, variable numbers of short consensus repeats (SCRs), an epidermal growth factor-like (EGF-like) domain and an N-terminal lectin domain. Their physiological ligands such as the P-selectin glycoprotein ligand-1 (PSGL-1) (McEver and Cummings 1997; McEver 2001) and E-selectin ligand-1 (ESL-1) (Steege et al. 1995) contain the common tetrasaccharide epitope sialyl Lewis^x (sLe^x, NeuNAcα2-3Galβ1-4(Fucα1-3)GlcNAc; Figure 1A) as part of O-glycosylated glycoproteins. The structural analysis of E- and P-selectin soaked with sLe^x revealed the central role of the L-fucose moiety, which coordinates the Ca²⁺ ion with its 3- and 4-hydroxyl groups (Somers et al. 2000).

In recent years, several case-control studies have suggested a correlation of a single-nucleotide polymorphism (SNP) in exon 4 of the E-selectin encoding gene with an early onset, increased susceptibility and severity of various, mainly cardiovascular and immune diseases. These include myocardial infarction (Yoshida et al. 2003; Stepień et al. 2011), severe recurrent venous thromboembolism (Jilma et al. 2005, 2006) and the formation of colon metastasis (Alessandro et al. 2007). The SNP causes a mutation of serine 128 to arginine (S128R) and is present in 10–15% of the Caucasian population (Wenzel et al. 1994). The mechanism by which the S128R mutation alters the physiological function of E-selectin at molecular level is still unclear, since the mutation is located in the EGF-like domain and not in the carbohydrate recognition domain (CRD) of the lectin. Based on the crystal structure of E-selectin (lectin and EGF-like domain), a direct or indirect involvement of the Ser128 in ligand binding is unlikely (Figure 1B and C) (Graves et al. 1994; Somers et al. 2000).

All previous reports on the binding properties of the S128R E-selectin mutant are summarized in Table I. First, Revelle et al. showed a 2- to 3-fold increased adherence of S128R E-selectin to HL60 cells compared with wt E-selectin (Revelle et al. 1996). Furthermore, only S128R E-selectin was capable of binding to heparin and to K562 cells, a cell line lacking fucosyltransferase VII required for sLe^x biosynthesis (Snapp et al. 1997). They also showed that the binding of heparin and K562 cell to the S128R E-selectin mutant was reversed by the sLe^x precursor 3'-sialyl-

¹To whom correspondence should be addressed: Tel: +41-61-267-1551; Fax: +41-61-267-1552; e-mail: beat.ernst@unibas.ch

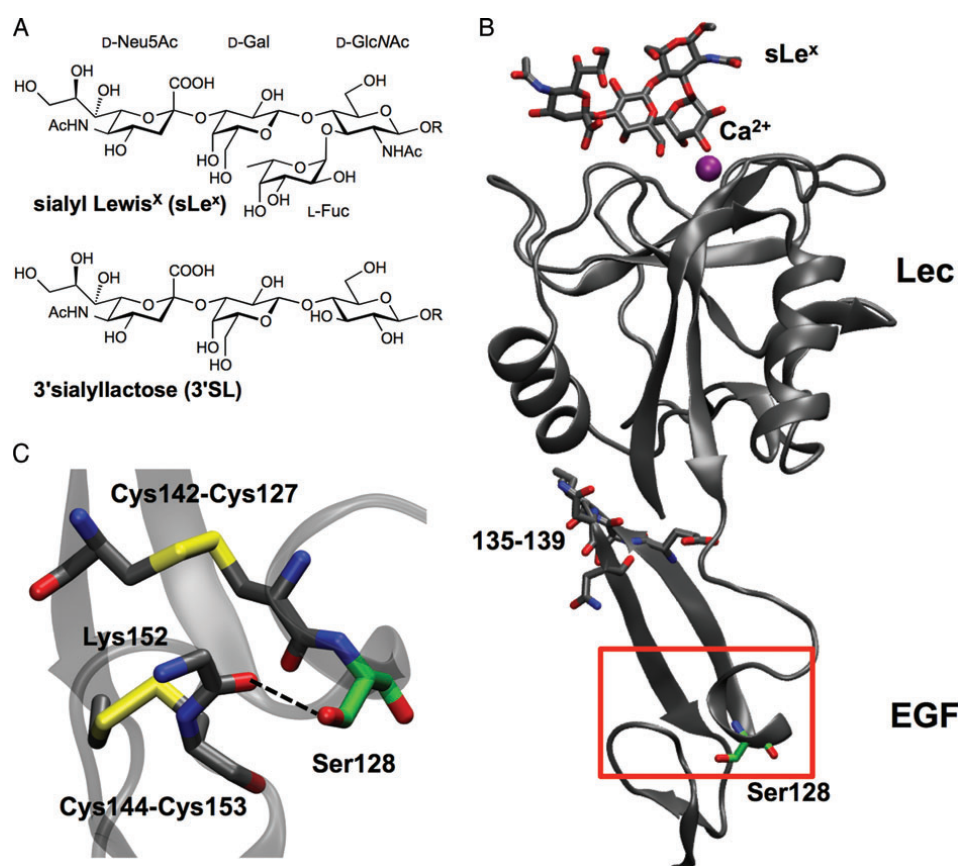


Fig. 1. (A) The natural ligand of E-selectin is sialyl Lewis^x (sLe^x). The proposed S128R E-selectin specific trisaccharide 3'sialyllactose lacks the L-fucose moiety. (B) The crystal structure of the wt E-selectin/sLe^x complex. The residues 135–139 form interactions with the lectin domain (Lec), EGF (epidermal growth factor like domain). (C) The side chain of the Ser128 residue forms a hydrogen bond to the backbone carbonyl of Lys152. This area is rigidified by two disulfide bonds (Cys144-Cys153 and Cys127-Cys142).

Table I. Summary of the reported binding properties of the S128R E-selectin mutant

	Revelle et al. (1996)	Wenzel et al. (1999)	Rao et al. (2002)	Yoshida et al. (2003)	This publication
Assay type	Protein–cell	Cell–cell	Cell–cell	Cell–cell	Protein–ligand/cell
Assay conditions	Static	Static	Flow	Flow	Static
wt and S128R	Soluble Lec-EGF-CR2/IgG	Full length on COS	Full length on CHO	Full length on HUVEC	Soluble Lec-EGF-CR2/IgG
E-selectin constructs					
HL60	Higher binding of S128R than wt; S128R not inhibited by sLe ^x or anti-sLe ^x -Ab	Lower binding of S128R than wt	Higher binding of S128R than wt	Higher binding of S128R than wt	Equal binding of wt and S128R
K562	S128R binding not inhibited by anti-sLe ^x -Ab	n.d.	Binding only of S128R	n.d.	No binding of S128R
3'SL/3'SLN	Inhibition of S128R binding to K562 and fetuin	n.d.	n.d.	n.d.	No binding of S128R
Heparin	Binding only of S128R	n.d.	No binding of S128R	n.d.	No binding of S128R
Bovine fetuin	Binding only of S128R	n.d.	n.d.	n.d.	No binding of S128R
sLe ^x /sLe ^a	n.d.	n.d.	n.d.	n.d.	Equal binding of wt and S128R

n.d., not determined.

N-acetylglucosamine (3'SLN, NeuNAcα2-3Galβ1-4GlcNAc) or 3'sialyllactose (3'SL, NeuNAcα2-3Galβ1-4Glc) (Figure 1A), indicating that the presence of the L-fucose moiety was no longer

required. Furthermore, [Revelle et al. \(1996\)](#) demonstrated that the binding of HL60 cells (which express the sLe^x epitope) was abolished with soluble sLe^x or an anti-sLe^x antibody for wt E-selectin,

but did not affect binding of S128R E-selectin. Finally, the S128R E-selectin mutant showed specific binding to bovine fetuin, a glycoprotein lacking fucosylated glycans (Green et al. 1988), further suggesting a shift in ligand specificity.

In contrast, Wenzel et al. (1999) reported cell-based assays under static conditions that showed diminished adhesion of HL60 cells to a COS cell monolayer bearing the S128R E-selectin mutant compared with the wt E-selectin counterpart.

Later, Rao et al. (2002) observed that only chinese hamster ovarian cells (CHO) cells displaying the S128R E-selectin mutant were able to tether K562 cells under flow conditions, an effect that was not observed with wt E-selectin. However, no inhibitory effect of heparin for S128R E-selectin as suggested by Revelle et al. (1996) was observed. Since neuraminidase or O-sialogly coprotein endopeptidase treatment of the K562 cells was not deleterious for their tethering toward S128R E-selectin expressing CHO cells, Neu5Ac- as well as O-linked glycoprotein-independent binding to the S128R mutant was implied.

Finally, Yoshida et al. (2003) reported increased attachment of HL60 cells toward human umbilical vein cells (HUVECs) displaying S128R E-selectin compared with those displaying wt E-selectin.

Since the S128R E-selectin polymorphism is correlated with numerous diseases, we aimed to re-evaluate its binding specificity at molecular level and additionally determine the potency of E-selectin antagonists which were originally designed for the wt E-selectin. We used a cell-free, target-based assay, which is routinely applied for screening of E-selectin antagonists (Weitz-Schmidt et al. 1996; Thoma et al. 1997). In addition, in a cell-based assay setup, we investigated the binding of HL60 and K562 cells under static and flow conditions.

Results

Characterization of wt and S128R E-selectin

For the affinity determination of three E-selectin antagonists and for further functional analyses, we recombinantly expressed a wt E-selectin construct containing the lectin domain, the EGF-like domain and the first two SCR domains and the corresponding S128R mutant as IgG chimera. The binding properties of the proteins were evaluated in a cell-free, target-based assay and additionally in a cell-based assay under static and flow conditions.

The proteins were expressed in CHO cells followed by functional purification with the functional monoclonal anti-E-selectin antibody 7A9 (see experimental part). The fact that the 7A9-antibody bound S128R E-selectin suggests correct folding of the protein. The sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) analysis showed a similar apparent molecular weight of 100 kDa, suggesting equal glycosylation of the wt and the mutant E-selectin (Figure 2A). The presence of the mutation was confirmed by DNA sequencing and at protein level by liquid chromatography-tandem mass spectrometry analysis (data not shown).

For the evaluation of protein activity, we determined the binding affinities of the natural carbohydrate ligands sLe^x and sLe^a to wt and S128R E-selectin as shown in Figure 2B. For wt E-selectin, the EC₅₀ values for biotinylated sLe^x- and sLe^a-polyacrylamide polymer (sLe^x-PAA, sLe^a-PAA) were 119.5 ± 17.1 ng/mL and 26.2 ± 1.3 ng/mL, respectively. The higher affinity of

sLe^a compared with sLe^x for recombinant E-selectin has previously been reported with a comparable assay (Weitz-Schmidt et al. 1996; Thoma et al. 1997). Similar affinities were also obtained with S128R E-selectin, with EC₅₀ values of 93.8 ± 13.1 ng/mL for sLe^x-PAA and 27.2 ± 1.2 ng/mL for sLe^a-PAA. Using concentrations of the biotinylated 3'SL-polyacrylamide polymer (3'SL-PAA) of up to 3 µg/mL, no binding was observed for both wt and S128R E-selectin, suggesting no or very low affinity. The evaluation of higher concentrations of 3'SL-PAA was not possible, as the background signal increased significantly.

sLe^x glycomimetics are efficacious for S128R E-selectin

The binding affinities of three previously reported E-selectin antagonists for wt and S128R E-selectin are shown in Figure 2C. In all the three of these sLe^x mimics, the Neu5Ac moiety has been exchanged with (S)-cyclohexyl lactic acid. For antagonist 1 (Norman et al. 1998), GlcNAc has been replaced with (1R,2R)-cyclohexane-1,2-diol and for antagonist 2 and 3 (Schwizer et al. 2012), with (1R,2R,3S)-3-methylcyclohexane-1,2-diol. For antagonist 3, the 2-hydroxyl group of D-galactose is additionally benzoylated. The half-maximal inhibitory concentration (IC₅₀) values of all three ligands were similar for the wt as well as S128R E-selectin. In agreement with previously reported results (Schwizer et al. 2012), the affinities of compounds 2 and 3 were approximately 4- and 14-fold higher than that of antagonist 1, respectively. The absolute IC₅₀ values obtained here are, however, weaker than previously reported due to the presence of only two SCR domains instead of six SCR domains, which is known to reduce E-selectin affinity (Li et al. 1994).

Evaluation of putative S128R E-selectin ligands

According to Revelle et al. (1996), bovine fetuin and heparin directly interact with S128R E-selectin, and 3'SLN interferes with the mutant binding to K562 cells. For the evaluation of their affinities (see experimental part) single concentrations of the carbohydrate ligands (3'SLN: 10 mM, bovine fetuin: 10 mg/mL, heparin: 10 mg/mL) were used to enable clear inhibition. As a result, the binding of sLe^x-PAA to wt and S128R E-selectin was not inhibited by any of the three substances (Figure 3A). In a control experiment, the assay was also performed with wt P-selectin, for which heparin is a known ligand (Nelson et al. 1993; Wang and Geng 2003). As shown in Figure 3A, sLe^x-PAA binding to wt P-selectin was completely abolished in the presence of heparin.

In addition to the target-based assays, further investigations of the binding properties of S128R E-selectin were performed with a cell-based assay under static and flow conditions. In the static assay, the adhesion of HL60 and K562 cells to wt and S128R E-selectin was determined. Revelle et al. (1996) reported that under static conditions the sLe^x-presenting HL60 cells adhere to both wt and mutant E-selectin, whereas the sLe^x-deficient K562 cells selectively bind S128R E-selectin. In our assay, fluorescently labeled HL60 and K562 cells were allowed to bind to E-selectin coated on microtiter plates (see Figure 3B). HL60 cells bind to both forms of E-selectin and since the addition of the functional anti-E-selectin antibody 7A9 inhibited binding entirely, both interactions are clearly E-selectin dependent. For the K562 cells, however, no binding to either form of E-selectin was observed, albeit the high number of applied cells. In the flow assay, adhesion of HL60

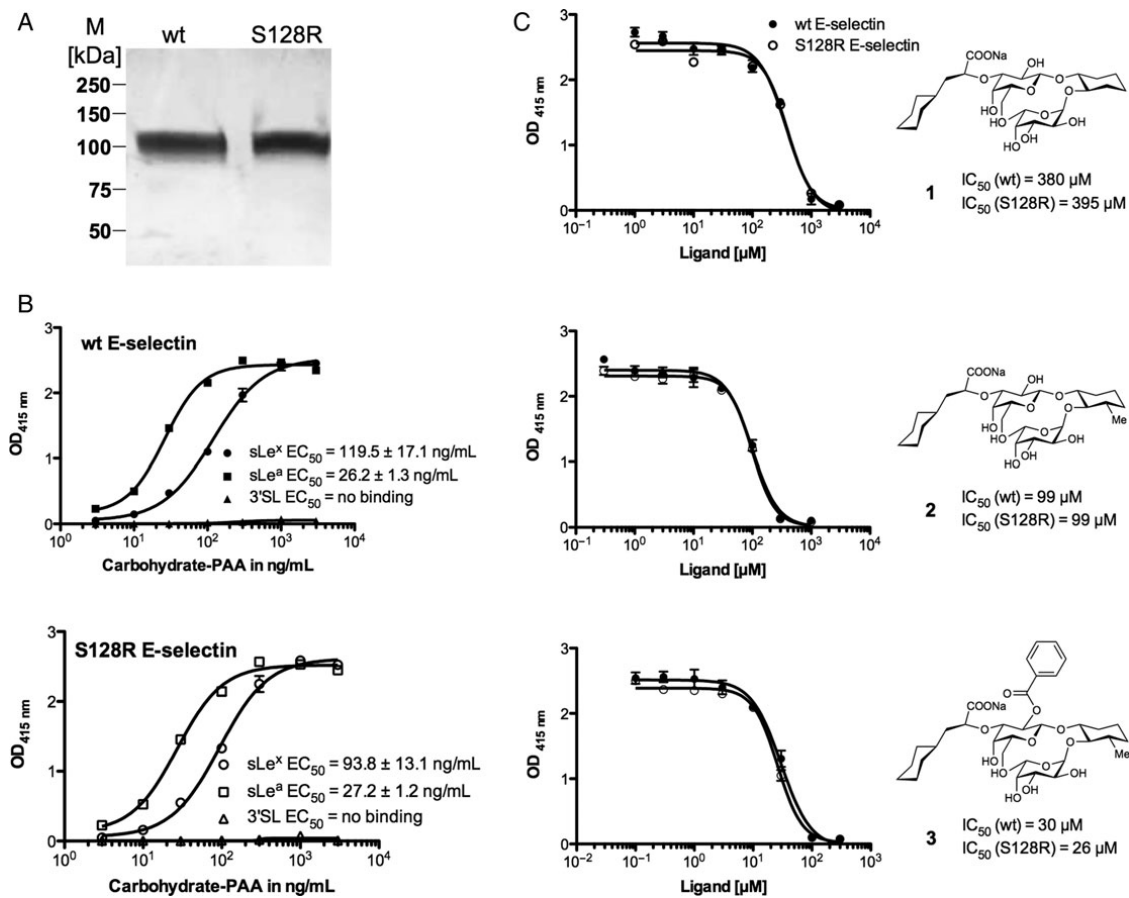


Fig. 2. (A) SDS-PAGE analysis of recombinant wt and S128R E-selectin purified by affinity chromatography on a 7A9-Sepharose column as described in the Materials and methods section. M, molecular weight marker. (B) Determination of the EC_{50} values for sLe^x, sLe^a and 3'SL-PAA with wt E-selectin (B, upper panel) and S128R mutant (B, lower panel). For this assay microtiter plates were coated overnight with either wt or S128R E-selectin (3 μ g/mL), blocked with BSA (2%) and incubated with increasing concentrations of sLe^x-PAA. The bound polymer was detected by colorimetric reaction at 415 nm. The assay was repeated three times, each in duplicate as described in the Materials and methods section. (C) Competitive binding assay for the determination of IC_{50} for antagonists 1-3 with wt and S128R E-selectin. The assay was performed as described in the Materials and methods section.

cells was studied at physiologic shear stress (Figure 3C). Again, the number of initially tethered and consequently rolling cells was similar for wt and S128R E-selectin. While at low shear stress (1 dyne/cm²) the highest number of cells tethered and rolled on wt and mutant E-selectin coated surfaces, decreased tethering was observed when higher shear forces were applied (2 and 3 dynes/cm²). At 5 dynes/cm², tethering was entirely abolished. Addition of EDTA completely abolished HL60 cells from binding to both forms of E-selectin, thus confirming the calcium-dependent binding (data not shown).

Glycan array screening

To identify potential novel ligands of S128R E-selectin, a glycan array screening was performed at the Core H facilities of the Consortium for Functional Glycomics. In this analysis, 465 glycan structures (version 4.1) were tested, including the tetrasaccharides sLe^a and sLe^x and derivatives thereof such as 3'SLN, Le^a and Le^x. For comparison, wt and S128R E-selectin were analyzed in parallel. Binding of the E-selectin constructs was detected with

Alexa Fluor[®] 488 labeled anti-IgG antibody and was quantified by measuring fluorescence intensities as shown in Figure 4.

The results clearly indicate that both proteins preferentially bind to glycan structures bearing sLe^a and sLe^x, thus confirming the result of the competitive binding assays. In addition, binding to 3-sulfated Lewis^a and Lewis^x was also observed for both E-selectin constructs. Finally, the S128R mutant did not exhibit any affinity for 3'SLN or any other non-fucosylated glycan. For both proteins, the ten best binding glycan structures contain a fucose moiety. This result underlines the importance of the fucose moiety for the binding of wt E-selectin and the corresponding mutant S128R E-selectin.

Discussion

A SNP of E-selectin in which the serine residue at position 128 is substituted by arginine was shown to correlate with various cardiovascular and immune diseases (Yoshida et al. 2003;

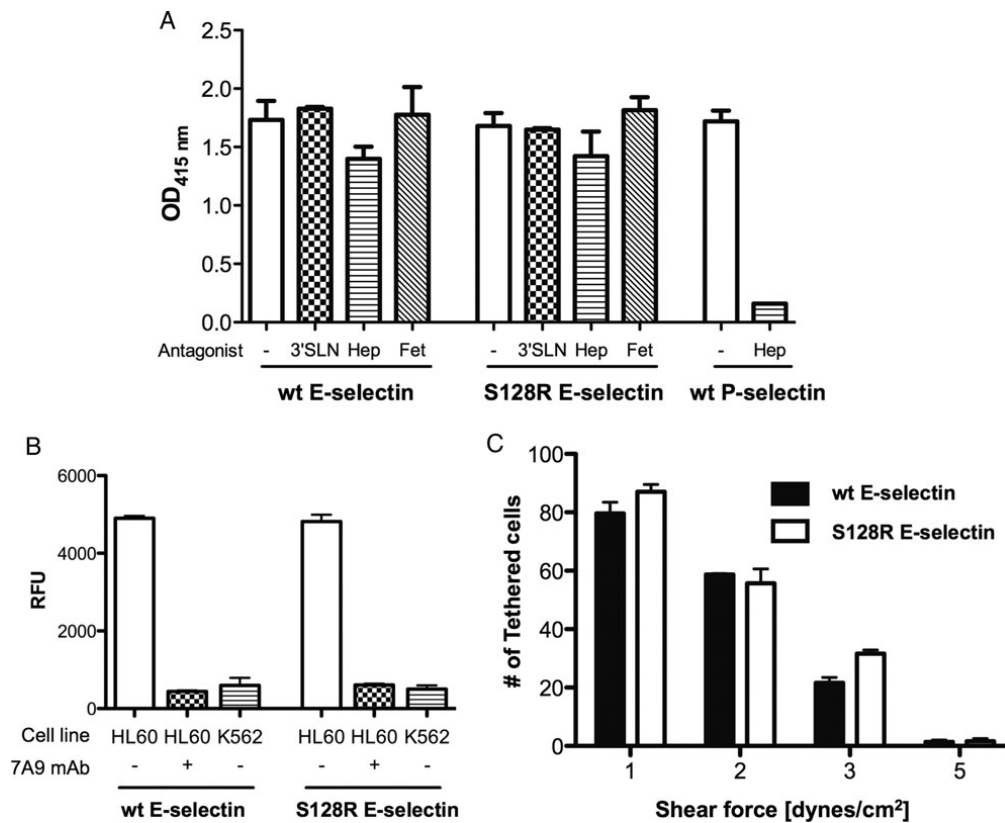


Fig. 3. Evaluation of the binding of 3'SLN (10 mM), heparin (Hep) (10 mg/mL) and bovine fetuin (Fet) (10 mg/mL) to wt E-selectin and S128R mutant (A). Microtiter plates were coated with either wt or S128R E-selectin (3 µg/mL), blocked with BSA (2%) and incubated with a fixed concentration of sLe^x-PAA (20 ng/mL), either in the presence or in the absence of Hep and Fet. As positive control wt P-selectin was co-incubated with sLe^x-PAA (20 ng/mL) and heparin (10 mg/mL). (B) Cell-based static binding assay with HL60 and K562 cell lines. Microtiter plates were coated with either wt or S128R E-selectin (10 µg/mL), blocked with BSA (2%) and incubated with the fluorescently labeled cells (2×10^5 cells/well), in the presence or absence of anti-E-selectin antibody 7A9 (20 µg/mL). Bound cells were measured by fluorescence measurements at 492 nm excitation and 517 nm emission wavelengths. The assay was repeated twice, each in duplicate as described in the Materials and methods section. (C) Cell-based flow binding assay with HL60 cells at physiological shear stress. Flow chambers were coated with protein A (50 µg/mL) and subsequently with wt or S128R E-selectin (50 µg/mL). The HL60 cells (1×10^6 cells/mL) were perfused at given shear stress (1–5 dynes/cm²). Cells were counted after 10 min at three points along the flow chamber. The error bars correspond to the standard error of the mean of the three points.

Jilma et al. 2006; Alessandro et al. 2007). The mutation is located in the EGF-like domain of E-selectin, which is necessary for, but not directly involved in, ligand binding (Erbe et al. 1992). Reville et al. (1996) reported that the regular physiological E-selectin ligand, the tetrasaccharide sLe^x was no longer required for S128R E-selectin to interact with HL60 and K562 cells.

In competition assays with K562 cells or heparin, 3'SL and 3'SLN were identified as ligands for S128R E-selectin. Furthermore, the glycoproteins heparin and bovine fetuin were shown to interact only with S128R E-selectin and not with wt E-selectin. Interestingly, the E-selectin R84A mutation, which is located adjacent to the carbohydrate-binding site, showed identical binding behavior as the S128R mutant. Binding of E-selectin-R84A to fetuin, however, was not observed in a previous publication, notably with the identical assay (Kogan et al. 1995). In contrast, Rao et al. (2002) reported that heparin did not inhibit S128R E-selectin binding to K562 cells. As stated above, the adverse effects of the S128R mutation were

investigated in a number of case-control studies for a variety of diseases, indicating that the loss of ligand specificity is possibly correlated with an increased leukocyte recruitment which in turn leads to pathophysiological situations.

Selectin antagonists have unequivocally been shown to interfere with excessive recruitment of leukocytes upon inflammation in vivo (Mulligan et al. 1993; Norman et al. 1998; Ridger et al. 2005). For the development of such antagonists, functional and structural information on the target protein is a prerequisite. For this purpose and due to the controversial reports on the binding properties of the S128R mutant, we decided to express the S128R mutant and the corresponding wt E-selectin as secreted IgG forms. Using cell-free target-based and cell-based binding assays, we assessed the binding properties of S128R E-selectin in comparison with the wild-type (wt). For wt and mutant E-selectin, no significant difference in the affinity for sLe^a and sLe^x was observed, indicating that the carbohydrate-binding site of the E-selectin mutant is not altered.

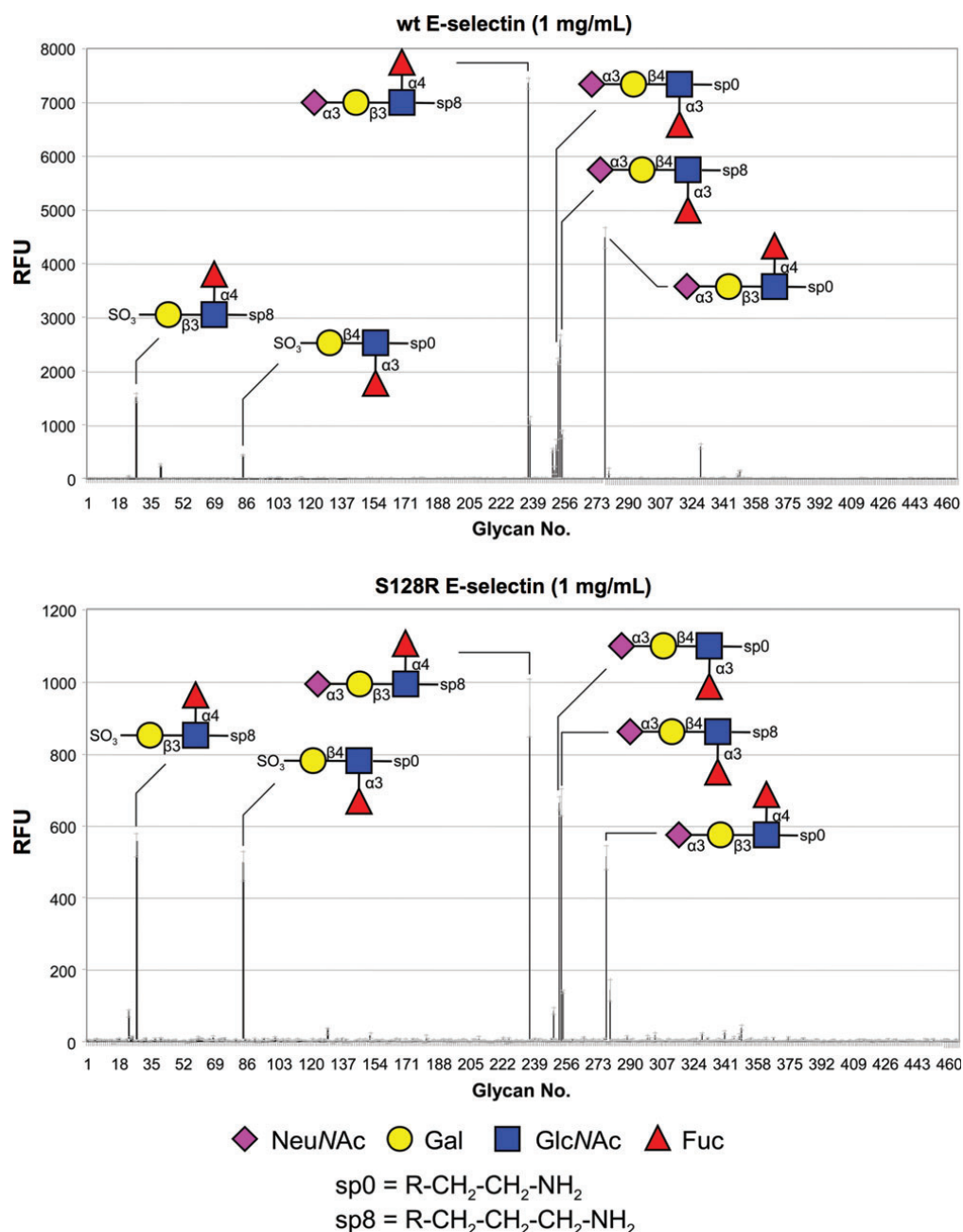


Fig. 4. Glycan array screening (version 4.1) of wt and S128R E-selectin. Binding was detected by a second incubation with Alexa Fluor[®] labeled anti-IgG antibody as described in the Materials and methods section. The values of the RFUs are given the average of 6 measurements of which the lowest and highest values were removed.

To investigate a potential treatment of patients carrying the S128R E-selectin polymorphism, binding affinities of three E-selectin antagonists were determined (Figure 2C). We demonstrate that regardless of the S128R mutation, binding affinities for all antagonists were similar when compared with wt E-selectin. This further confirmed that carbohydrate-binding

site of S128R E-selectin was intact and suggests a possible application of such sLe^x mimics in the treatment of affected patients. Regarding the specificity of S128R E-selectin, direct binding of the 3'SL polymer was not observed even at high concentrations. To further evaluate the promoted specific ligands for S128R E-selectin, inhibition assays with sLe^x

polymer in competition with the 3'SLN trisaccharide, heparin and bovine fetuin were performed. Even at high concentrations, none of these three compounds showed inhibition of the sLe^x polymer binding to S128R E-selectin (Figure 3A). These results suggest similar binding phenotypes of wt and S128R E-selectin. Finally, to investigate the binding of S128R E-selectin to leukocytes, we performed cell-based assays (Figure 3B and C). The sLe^x bearing HL60 cell line bound wt and S128R E-selectin to the same degree under static and physiological flow conditions and in an E-selectin dependent manner. Again, no difference in binding specificity was observed for S128R E-selectin. Furthermore, unlike previously reported, no specific binding of K562 cells was measured. Therefore, no interaction between S128R E-selectin and an unknown carbohydrate or non-carbohydrate ligand on K562 occurred. Finally, to characterize the binding phenotype with another cell-free assay, a glycan array screening was performed at the Consortium for Functional Glycomics (Figure 4). The binding of wt and S128R E-selectin to 465 different mammalian glycan structures was evaluated. The results confirmed the similar binding profiles for wt and S128R E-selectin. The 10 best binding glycan structures all contained a fucose moiety. The RFU values of glycans lacking the fucose moiety such as 3'SLN showed signals on background level for both proteins. Additionally, no specific ligand for the S128R E-selectin mutant was detected, and thus, its binding properties do not differ from those of the wt. As in the EC₅₀ determinations of sLe^a and sLe^x, an augmented binding of sLe^a over sLe^x was also observed which is reflected by the higher RFU values for sLe^a. Although high E-selectin concentrations (1 mg/mL) were applied for the glycan array, the resulting RFU values were rather low. This was most likely due to the low affinity of E-selectin for its tetrasaccharide ligands (Schwizer et al. 2012). The reason why S128R E-selectin showed eightfold lower maximal RFU values compared with wt E-selectin is not clear, since in the polymer-based activity assays, the affinity (EC₅₀) of both proteins for sLe^x- and sLe^a-PAA were comparable (equal maximal OD_{415nm} values). However, glycan array screenings provide for qualitative and not quantitative measurements.

Our results suggest similar binding specificities of wt and S128R E-selectin under static and flow conditions. This is to be expected from a visual inspection of the E-selectin crystal structure in complex with sLe^x (Somers et al. 2000). It was shown that only the amino acids 135–139 from the EGF-like domain form a direct contact with the lectin domain by hydrogen bonding and van der Waals interactions (Graves et al. 1994). The distance between the Ser128 C α atom and Ca²⁺ in the binding site is 43 Å. The Ser128 side chain forms a hydrogen-bond with the backbone carbonyl of Lys152, an interaction found in all selectin crystal structures. This hydrogen bond would be lost in case of the S128R mutation. The fact that the amino acid sequence Ser126-Cys127-Ser128 is conserved in all the three selectins and throughout several species implies that this region is indeed important for the proper function of the selectins. However, the region surrounding the mutation is rigidified by three disulfide bonds (Cys122–Cys133, Cys127–Cys142 and Cys144–Cys153) within a 10 Å radius of Ser128 C α atom (Figure 1C).

In conclusion, as the S128R E-selectin displayed unaltered binding specificity under static and flow conditions as wt E-selectin, we conclude that glycomimetic E-selectin antagonists

are also effective antagonists of the mutant. According to our studies, the reported correlation of a SNP in the E-selectin encoding gene with an increased susceptibility and severity of various, mainly cardiovascular and immune, diseases is therefore not related to altered binding properties of wt and mutant E-selectin. Further investigations, e.g., structural and biophysical protein characterization, as well as functional in vivo assays are necessary to shed light on possible therapeutic opportunities based on the S128R E-selectin mutant.

Materials and methods

Materials and reagents

The E-selectin primers lec-fw (5'-GGCC GAATTC GTG GTC TTA CAA CAC CTC CAC GGA A-3') and cr2_rev (5'-GGCC GATATC GAA GCT TTA CAC GTT GGC TTC TCG TT-3') containing the restriction sites EcoRI and EcoRV (underlined), respectively, and the mutation primers egf_s128r_fw (5'-AAT ACA TCC TGC AGA GGC CAC GGT-3') and egf_s128r_rev (5'-ACC GTG GCC TCT GCA GGA TGT ATT-3') were purchased from Microsynth (Balgach, Switzerland). The restriction enzymes EcoRI and EcoRV and the glycosidase PNGaseF were obtained from New England BioLabs (Allschwil, Switzerland). The pFUSE-IgG2-Fc2 expression vector and the antibiotic ZeocinTM were purchased from Invivogen (Toulouse, France). Tissue culture flasks and MaxiSorpTM 96-well microtiter plates were from Nunc (Roskilde, Denmark). Ham's Nutrient Mixture F-12 medium, Roswell Park Memorial Institute (RPMI-15) medium, fetal calf serum (FCS) and the horseradish peroxidase substrate 2,2'-azino-di(3-ethylbenzthiazoline-6-sulfonic acid) (ABTS) were purchased from Invitrogen (Lucerne, Switzerland). Penicillin (10,000 U/mL)/streptomycin (10 mg/mL), cyanogen bromide-activated-Sepharose[®] 4B, bovine fetuin and heparin sodium salt were obtained from Sigma-Aldrich (Basel, Switzerland). The FuGENE[®] HD transfection reagent and the streptavidin-peroxidase (streptavidin-POD) were purchased from Roche Applied Science (Rotkreuz, Switzerland). Amicon[®] ultrafiltration tubes (50 kDa cut-off) were obtained from Millipore (Zug, Switzerland). The ion exchange column UnoQ6 was purchased from Bio-Rad (Reinach BL, Switzerland). Protein A-Sepharose[®] was obtained from BioVision (Mountain View, CA). Protein G-Sepharose[®] was purchased from Amersham Pharmacia (GE Healthcare, Piscataway, NJ). 3-SL, sLe^x- and sLe^a-polyacrylamide (PAA)-biotin were from GlycoTech (Gaithersburg, MD). The Alexa Fluor[®] 488 labeled anti-human-IgG(Fc) antibody and the Microscanarray XL4000 scanner were from Invitrogen (Carlsbad, CA) and PerkinElmer Life Sciences (Waltham, MA), respectively. Trypsin sequencing grade was obtained from Promega (Dübendorf, Switzerland). The trapping 300SB C-18 column (0.3 × 50 mm) was purchased from Agilent Technologies (Basel, Switzerland), the Magic 300 Å C18 reverse-phase material (5 µm particle size) from Michrom Bioresources, Inc. (Auburn, CA), the Orbitrap FT hybrid instrument from Thermo Finnigan (San Jose, CA) and the Rheos 2200 pump from (Flux Instruments, Basel, Switzerland). The cell line HL60 was purchased from LGC Standards (Molsheim, France) and K562 cells from the German Collection of Microorganisms and Cell Cultures (DSMZ, Braunschweig, Germany). The BioFluxTM system was purchased from Fluxion (San Francisco, CA).

Molecular cloning

Genomic DNA from wild-type human E-selectin [Lec-EGF-CR6/hIgG1(Fc)] expressing CHO-K1 cells (Scheffler et al. 1995) was used for polymerase chain reaction (PCR) amplification of the cDNA encoding for the human E-selectin lectin domain (bp 1–360), the EGF-like domain (bp 361–471) and the first two consensus repeat domains (bp 472–840). The forward primer *lec_fw* and the reverse primer *cr2_rev* containing EcoRI and EcoRV restriction sites were used. The restricted insert was ligated into the corresponding cloning site of the pFUSE-hIgG2-Fc2 expression vector. The resulting construct containing the N-terminal interleukine-2 secretion signal and the Fc part of human IgG2 at C-terminus was transformed into chemocompetent *Escherichia coli* DH5 α . After single clone selection and DNA-minipreparation, the correctness of the construct was confirmed by DNA double-strand sequencing (Microsynth, Balgach, Switzerland).

Site-directed mutagenesis was employed for the thymine-to-adenine substitution at cDNA position 384, shifting the expression from a serine to an arginine at amino acid position 128. The standard overlap extension PCR method was used (Ho et al. 1989). In a first step, two overlapping DNA fragments were generated separately, using wild-type E-selectin cDNA as template. The first fragment was amplified with the *lec_fw* primer and the internal reverse primer *egf_s128r_rev* and the second fragment was amplified using *cr2_rev* and the internal forward primer *egf_s128r_fw*. The two internal primers contained a mismatch for the site-directed base substitution. In a second step, both overlapping DNA fragments were elongated to the full-length insert, containing the single point mutation. The cloning into the pFUSE-hIgG2-Fc2 expression vector and sequence verification were performed as described above for the wt construct.

Cell transfection and expression of wt E-selectin, wt P-selectin and S128R E-selectin

CHO-K1 cells (American Type Culture Collection No. CCL-61™) were cultivated in tissue culture flasks in Ham's Nutrient Mixture F-12 medium supplemented with 2 mM L-glutamine, 10% (v/v) FCS, 100 U/mL penicillin and 100 μ g/mL streptomycin. The cells were transfected with either the wt or the S128R construct using the FuGENE® HD transfection reagent according to the instructions of the supplier (Roche). The selection of stably transfected CHO-K1 cells was achieved by treatment with 500 μ g/mL Zeocin™ and single clones were obtained by limiting dilution. Culture medium containing the secreted E-selectin [Lec-EGF-CR2/hIgG2(Fc)] chimera was harvested weekly and sterile filtered (0.22 μ m). The medium was used immediately for protein purification or stored at –20°C.

For control experiments P-selectin [Lec-EGF-CR6/hIgG1(Fc)] wt expressed in CHO-K1 cells was used (Fritz et al. 1998).

Expression and purification of anti-E-selectin antibody 7A9

Mouse hybridoma cells expressing the mouse monoclonal anti-E-selectin antibody 7A9 (American Type Culture Collection No. HB-10135™) were cultivated in RPMI-15 medium supplemented with 15% (v/v) FCS, 100 U/mL penicillin and 100 μ g/mL streptomycin. Antibodies were purified from the culture medium by protein G-Sepharose®. The column was attached to a fast protein liquid chromatography apparatus (Bio-Rad, Reinach BL,

Switzerland) and equilibrated with loading buffer (50 mM Tris-HCl (pH 7.6), 150 mM NaCl and 0.05% v/v Tween-20). The culture medium was filtrated (0.22 μ m) diluted with loading buffer and then loaded onto the protein G-Sepharose® column followed by two washing steps, first with loading buffer and then with washing buffer (5 mM NH₄OAc, pH 5.0). The protein was finally eluted with elution buffer (0.5 M NH₄OAc, pH 3.4) and the eluate was immediately neutralized to pH 7–7.5 with a 2.5 M Tris solution. The protein concentration was determined by high-performance liquid chromatography (HPLC) (Mesch et al. 2012). Approximately 5 mg of purified 7A9 antibody were coupled to 1 mL of cyanogen bromide-activated-Sepharose® 4B according to the manufacturer's protocol.

Purification of wt E-selectin, wt P-selectin and S128R E-selectin

The culture medium containing E-selectin was diluted (1:1, v/v) with loading buffer, filtrated (0.22 μ m) and purified with the 7A9-Sepharose® column using the same protocol and buffers as for the purification of the 7A9 antibody (see above). The purified protein was concentrated by ultrafiltration (50 kDa cut-off).

P-selectin was purified in two steps. The first step was performed on a protein A-Sepharose® column using the same procedure and buffers as described above. The second step was performed with anion exchange chromatography using a UnoQ6 column and the buffer A (20 mM Tris-HCl, pH 7.4) for column equilibration and loading and buffer B (20 mM Tris-HCl, pH 7.4, 1 M NaCl) for protein elution.

Protein purity was confirmed by SDS-PAGE analysis followed by Coomassie Brilliant Blue G-250 staining. The protein concentration was determined by HPLC as previously reported (Mesch et al. 2012).

Activity binding assay

Biotinylated sLe^x-, sLe^a- and 3'SL-polyacrylamide polymers (sLe^x-PAA, sLe^a-PAA and 3'SL-PAA, purchased from Glyco-Tech, Gaithersburg, MD 20879) were coupled to streptavidin- β -peroxidase (POD) by incubation of 20 μ L (100 μ g/mL) polymer, 80 μ L (500 U/mL streptavidin-POD conjugate), 10 μ L FCS and 80 μ L assay buffer (20 mM HEPES pH 7.4, 150 mM NaCl, and 10 mM CaCl₂) for 2 h at 37°C. The complexes were stable for several weeks when stored at 4°C.

MaxiSorp™ 96-well microtiter plates were coated with 100 μ L/well of wt or S128R E-selectin at 3 μ g/mL in assay buffer overnight at 4°C in a humidified chamber. Control wells were coated with assay buffer without protein for subsequent background signal subtraction. The wells were washed three times with 150 μ L/well of assay buffer and blocked with 150 μ L/well of 3% w/v bovine serum albumin (BSA) in assay buffer for 3 h at 4°C. The wells were then washed three times with 150 μ L/well assay buffer before adding 100 μ L/well of a serial dilution (0–3 μ g/mL) of a sLe^x-, sLe^a- or 3'SL-PAA solution. Each concentration was applied in duplicate. After incubation for 3 h at 25°C and 250 rpm, the wells were carefully washed with 150 μ L/well assay buffer. Carbohydrate-PAA binding was detected by addition of 100 μ L/well of ABTS-substrate. The colorimetric reaction was allowed to develop for 10 min and then stopped by the addition of 2% aqueous oxalic acid before the optical density (OD) was measured at 415 nm on a microplate reader. The

OD_{415 nm} values were background subtracted and the EC₅₀ values were calculated using the Prism software (GraphPad Software, Inc., La Jolla). The assays were repeated three times on separate days. The EC₅₀ value defines the glycopolymer concentration corresponding to 50% of the maximum binding to the protein.

Competitive binding assay

For the evaluation of the proposed ligands for S128R E-selectin and the affinity measurements of the sLe^x mimics, polyacrylamide polymer-based competitive binding assay was performed. The trisaccharide 3'-sialyl-N-acetylglucosamine (3'SLN) was synthesized in-house as reported in the Supplementary data. Antagonists **1-3** were prepared as previously described (Schwitzer et al. 2012). MaxiSorp™ 96-well microtiter plates were coated with a solution of E- or P-selectin (3 µg/mL, 100 µL/well) and blocked for 3 h with 150 µL/well of BSA (3%, w/v). The wells were subsequently co-incubated with 50 µL of sLe^x-PAA (40 ng/mL) and 50 µL of 3'SLN (10 mM final concentration), bovine fetuin (10 mg/mL) or heparin sodium salt (10 mg/mL). Inhibition of sLe^x-PAA binding to wt P-selectin by heparin sodium salt served as positive control. For compounds **1-3**, 25 µL of serial dilutions were mixed with 25 µL of sLe^a-PAA polymer (40 ng/mL). All other assay steps were performed similar to the activity binding assay. For 3'SLN, the assay was repeated twice on separate days and for bovine fetuin and heparin three times on separate days. The sLe^x mimics were tested a single time in duplicate. IC₅₀ values were calculated using the Prism software. The IC₅₀ defines the molar concentration of the test compound that reduces the maximal specific binding of sLe^{x/a}-PAA to E/P-selectin by 50%.

Cell-based static assay

The protein coating, washing and blocking steps of the cell-based binding assay were performed similar to the activity binding assay, with an increase in wt or S128R E-selectin concentration to 10 µg/mL. HL60 and K562 cells were cultivated in suspension in RPMI 1640 culture medium supplemented with 10% FCS, 2 mM L-glutamine, 10% (v/v) FCS, 100 U/mL penicillin and 100 µg/mL streptomycin. The cell suspension was harvested at 5×10^5 – 1×10^6 cells/mL by centrifugation at 1500 rpm and 4°C. The cells were washed twice with PBS, resuspended to 1×10^6 cells/mL and labeled with 5 µM carboxyfluorescein succinimidyl ester) for 8 min at room temperature. The reaction was stopped with FCS (2%, v/v) and the cells were washed twice with assay buffer containing FCS (2%, v/v). Final concentrations of 2×10^5 cells/well and 20 µg/mL 7A9 antibody were used. After incubation for 1 h at 37°C and 250 rpm, the wells were washed three times with assay buffer. Fluorescent intensities were measured at 492 nm excitation and 517 nm emission with a fluorescence plate reader. The cell-based assays were performed in duplicate and repeated twice on separate days. Cell viability was >95% at all times as judged by trypan blue staining.

Cell-based flow assay

Flow chamber assays were performed with a BioFlux™ system. The flow chambers were coated over night at 4°C with 50 µg/mL soluble protein A in bicarbonate buffer (pH 9.6).

After washing with assay buffer, the chambers were perfused with wt or S128R E-selectin (50 µg/mL) and incubated for 2 h at 4°C. The chambers were subsequently blocked with 2% w/v BSA in assay buffer. HL60 cells were prepared as described above, resuspended to 1×10^6 cells/mL and perfused at physiological shear forces between 1 and 5 dynes/cm². Rolling cells were determined at three points along the chamber after 10 min perfusion with a Nikon inverted-stage phase-contrast microscope with 20× magnification. Shutter speed was set according to shear stress for correct distinction of adherent and free flowing cells.

Glycan array screening

The glycan array screening (version 4.1) was performed at the Core H facilities of the Consortium for Functional Glycomics as previously described (Blixt et al. 2004). Briefly, the wt and S128R E-selectin samples were diluted to 1 mg/mL in Tris-buffer [20 mM Tris-HCl (pH 7.4), 150 mM NaCl, 2 mM CaCl₂, 2 mM MgCl₂] supplemented with 0.05% Tween 20 and 1% BSA. An aliquot of 70 µL was applied to the slides containing the covalently linked glycans, each glycan in replicate of six. The slides were coverslipped and incubated in a humidified chamber protected from light for 1 h at room temperature and subsequently incubated with Alexa Fluor® 488 labeled anti-human-IgG(Fc) antibody at a concentration of 5 µg/mL in Tris-buffer. After washing and drying the slides, the fluorescence was measured with the Microscanarray XL4000 scanner. The relative fluorescence units (RFUs) were then calculated as the average of the six replicates of each glycan structure, of which the lowest and the highest values were ignored.

Supplementary data

Supplementary data for this article is available online at <http://glycob.oxfordjournals.org/>.

Funding

We thank the Swiss National Science Foundation for the financial support (200020_129935/1).

Acknowledgements

We acknowledge the Protein-Glycan Interaction Core (H) of the Consortium for Functional Glycomics funded by the NIGMS GM62116 for the glycan array analysis.

Conflict of interest

None declared

Abbreviations

3'SL-PAA, biotinylated 3'SL-polyacrylamide polymer; CRD, carbohydrate recognition domain; EGF-like, epidermal growth factor-like; HL-60, human promyelocytic leukemia cells; K562, human immortalized myelogenous leukemia cell line; Lec, lectin domain; PAA, polyacrylamide; POD, β-peroxidase;

PSGL-1, P-selectin glycoprotein ligand-1; RFU, relative fluorescent unit; SCR, short consensus repeat; sLe^a-PAA, biotinylated sLe^a-polyacrylamide polymer; sLe^x-PAA, biotinylated sLe^x-polyacrylamide polymer; SNP, single-nucleotide polymorphism, wt, wild type.

References

- Alessandro R, Seidita G, Flugy AM, Damiani F, Russo A, Corrado C, Colomba P, Gullotti L, Buettner R, Bruno L, et al. 2007. Role of S128R polymorphism of E-selectin in colon metastasis formation. *Int J Cancer*. 121:528–535.
- Blixt O, Head S, Mondala T, Scanlan C, Huflejt ME, Alvarez R, Bryan MC, Fazio F, Calarese D, Stevens J, et al. 2004. Printed covalent glycan array for ligand profiling of diverse glycan binding proteins. *Proc Natl Acad Sci USA*. 101:17033–17038.
- Erbe DV, Wolitzky BA, Presta LG, Norton CR, Ramos RJ, Burns DK, Rumberger JM, Rao BNN, Foxall C, Brandley BK, et al. 1992. Identification of an E-selectin region critical for carbohydrate recognition and cell-adhesion. *J Cell Biol*. 119:215–227.
- Ernst B, Magnani JL. 2009. From carbohydrate leads to glycomimetic drugs. *Nat Rev Drug Discov*. 8:661–677.
- Fritz J, Katopodis AG, Kolbinger F, Anselmetti D. 1998. Force-mediated kinetics of single P-selectin ligand complexes observed by atomic force microscopy. *Proc Natl Acad Sci USA*. 95:12283–12288.
- Graves BJ, Crowther RL, Chandran C, Rumberger JM, Li S, Huang KS, Presky DH, Familletti PC, Wolitzky BA, Burns DK. 1994. Insight into E-selectin/ligand interaction from the crystal-structure and mutagenesis of the lec/EGF domains. *Nature*. 367:532–538.
- Green ED, Adelt G, Baenziger JU, Wilson S, Vanhalbeek H. 1988. The asparagine-linked oligosaccharides on bovine fetuin. Structural analysis of N-glycanase-related oligosaccharides by 500-megahertz H-1-NMR spectroscopy. *J Biol Chem*. 263:18253–18268.
- Ho SN, Hunt HD, Horton RM, Pullen JK, Pease LR. 1989. Site-directed mutagenesis by overlap extension using the polymerase chain-reaction. *Gene*. 77:51–59.
- Jilma B, Kovar FM, Hron G, Endler G, Marsik CL, Eichinger S, Kyrle PA. 2006. Homozygosity in the single nucleotide polymorphism Ser128Arg in the E-selectin gene associated with recurrent venous thromboembolism. *Arch Intern Med*. 166:1655–1659.
- Jilma B, Marsik C, Kovar F, Wagner OF, Jilma-Stohlawetz P, Endler G. 2005. The single nucleotide polymorphism Ser128Arg in the E-selectin gene is associated with enhanced coagulation during human endotoxemia. *Blood*. 105:2380–2383.
- Kansas GS. 1996. Selectins and their ligands: Current concepts and controversies. *Blood*. 88:3259–3287.
- Kogan TP, Revelle BM, Tapp S, Scott D, Beck PJ. 1995. A single amino acid residue can determine the ligand specificity of E-selectin. *J Biol Chem*. 270:14047–14055.
- Li SH, Burns DK, Rumberger JM, Presky DH, Wilkinson VL, Anostario M, Jr, Wolitzky BA, Norton CR, Familletti PC, Kim KJ, et al. 1994. Consensus repeat domains of E-selectin enhance ligand binding. *J Biol Chem*. 269:4431–4437.
- McEver RP. 2001. Adhesive interactions of leukocytes, platelets, and the vessel wall during hemostasis and inflammation. *Thromb Haemost*. 86:746–756.
- McEver RP, Cummings RD. 1997. Role of PSGL-1 binding to selectins in leukocyte recruitment. *J Clin Invest*. 100:485–492.
- Mesch S, Lemme K, Wittwer M, Koliwer-Brandl H, Schwardt O, Kelm S, Ernst B. 2012. From a library of MAG antagonists to nanomolar CD22 ligands. *ChemMedChem*. 7:134–143.
- Mulligan MS, Paulson JC, Defrees S, Zheng ZL, Lowe JB, Ward PA. 1993. Protective effects of oligosaccharides in P-selectin-dependent lung injury. *Nature*. 364:149–151.
- Nelson RM, Cecconi O, Roberts WG, Aruffo A, Linhardt RJ, Bevilacqua MP. 1993. Heparin oligosaccharides bind L-selectin and P-selectin and inhibit acute-inflammation. *Blood*. 82:3253–3258.
- Norman KE, Anderson GP, Kolb HC, Ley K, Ernst B. 1998. Sialyl Lewis(x) (sLe(x)) and an sLe(x) mimetic, CGP69669A, disrupt E-selectin-dependent leukocyte rolling *in vivo*. *Blood*. 91:475–483.
- Rao RM, Clarke JL, Ortlepp S, Robinson MK, Landis RC, Haskard DO. 2002. The S128R polymorphism of E-selectin mediates neuraminidase-resistant tethering of myeloid cells under shear flow. *Eur J Immunol*. 32:251–260.
- Revelle BM, Scott D, Beck PJ. 1996. Single amino acid residues in the E- and P-selectin epidermal growth factor domains can determine carbohydrate binding specificity. *J Biol Chem*. 271:16160–16170.
- Ridger VC, Hellewell PG, Norman KE. 2005. L- and P-selectins collaborate to support leukocyte rolling *in vivo* when high-affinity P-selectin-P-selectin glycoprotein ligand-1 interaction is inhibited. *Am J Pathol*. 166:945–952.
- Scheffler K, Ernst B, Katopodis A, Magnani JL, Wang WT, Weisemann R, Peters T. 1995. Determination of the bioactive conformation of the carbohydrate ligand in the E-selectin sialyl Lewis(x) complex. *Angew Chem Int Ed Engl*. 34:1841–1844.
- Schwizer D, Patton JT, Cutting B, Smiesko M, Wagner B, Kato A, Weckerle C, Binder FP, Rabbani S, Schwardt O, et al. 2012. Pre-organization of the core structure of E-selectin antagonists. *Chemistry*. 18:1342–1351.
- Snapp KR, Wagers AJ, Craig R, Stoolman LM, Kansas GS. 1997. P-selectin glycoprotein ligand-1 is essential for adhesion to P-selectin but not E-selectin in stably transfected hematopoietic cell lines. *Blood*. 89:896–901.
- Somers WS, Tang J, Shaw GD, Camphausen RT. 2000. Insights into the molecular basis of leukocyte tethering and rolling revealed by structures of P- and E-selectin bound to SLe(X) and PSGL-1. *Cell*. 103:467–479.
- Steegmaier M, Levinovitz A, Isenmann S, Borges E, Lenter M, Kocher HP, Kleuser B, Vestweber D. 1995. The E-selectin-ligand ESL-1 is a variant of a receptor for fibroblast growth factor. *Nature*. 373:615–620.
- Stepień E, Krawczyk S, Kapelak B, Sobczyński R, Stoliński J, Wypasek E, Undas A, Sadowski J. 2011. Effect of the E-selectin Gene Polymorphism (S149R) on Platelet Activation and Adverse Events After Coronary Artery Surgery. *Arch Med Res*. 42:375–381.
- Thoma G, Magnani JL, Ohrlein R, Ernst B, Schwarzenbach F, Duthaler RO. 1997. Synthesis of oligosaccharide-polylysine conjugates: A well characterized sialyl Lewis(a) polymer for ELISA. *J Am Chem Soc*. 119:7414–7415.
- Wang JG, Geng JG. 2003. Affinity and kinetics of P-selectin binding to heparin. *Thromb Haemost*. 90:309–316.
- Weitz-Schmidt G, Stokmaier D, Scheel G, Nifantev NE, Tuzikov AB, Bovin NV. 1996. An E-selectin binding assay based on a polyacrylamide-type glycoconjugate. *Anal Biochem*. 238:184–190.
- Wenzel K, Hanke R, Speer A. 1994. Polymorphism in the human E-selectin gene detected by PCR-SSCP. *Hum Genet*. 94:452–453.
- Wenzel K, Stahn R, Speer A, Denner K, Glaser C, Affeldt M, Moobed M, Scheer A, Baumann G, Felix SB. 1999. Functional characterization of atherosclerosis-associated Ser128Arg and Leu554Phe E-selectin mutations. *Biol Chem*. 380:661–667.
- Yoshida M, Takano Y, Sasaoka T, Izumi T, Kimura A. 2003. E-selectin polymorphism associated with myocardial infarction causes enhanced leukocyte-endothelial interactions under flow conditions. *Arterioscler Thromb Vasc Biol*. 23:783–788.

Implications of the E-selectin S128R Mutation on Drug Discovery

Supplementary Data

Roland C. Preston¹, Said Rabbani¹, Florian P.C. Binder¹, Suzette Moes², John L. Magnani³, Beat Ernst¹

¹ Institute of Molecular Pharmacy, Pharmacenter, University of Basel, 4056 Basel, Switzerland

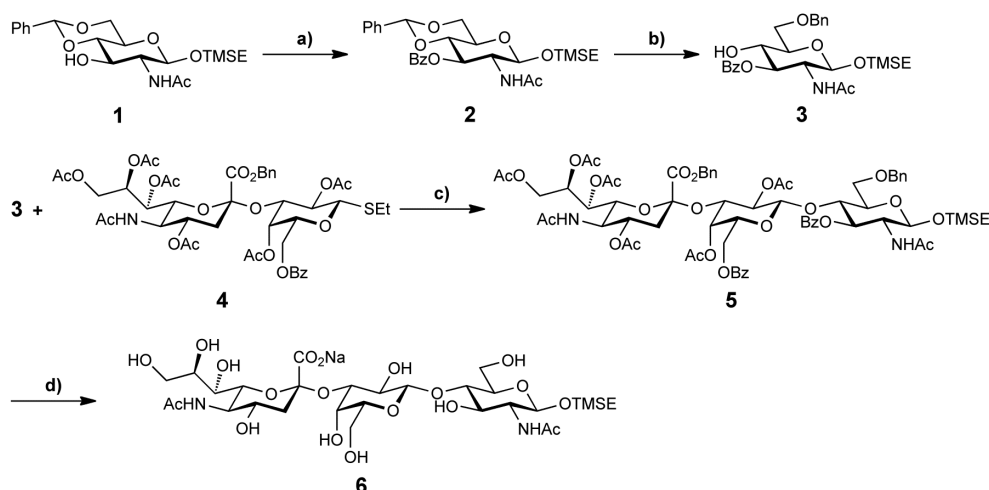
² Department of Biochemistry, Biocenter, University of Basel, 4056 Basel, Switzerland

³ GlycoMimetics, Inc., Gaithersburg, Maryland 20879, USA

Content

1. Synthesis of the trisaccharide 3'-sialyl-*N*-acetylactosamine (3'SLN, **6**)
2. LC-MS/MS analysis of E-selectin

1. Synthesis of the trisaccharide 3'-sialyl-*N*-acetylactosamine (3'SLN, **6**)



Scheme S1. a) BzCl, DMAP, pyridine, r.t., 4 h, 40%; b) Me₃N·BH₃, AlCl₃, THF, H₂O, r.t., 5 h, 79%; c) DMTST, CH₂Cl₂, MS 4 Å, r.t., 5 d, 39%; d) (i) H₂, Pd(OH)₂/C, dioxane, H₂O, r.t., 2 d; (ii) NaOH, MeOH, H₂O, r.t., 20 h, 77%.

General Methods: NMR spectra were recorded on a Bruker Avance DMX-500 (500 MHz) spectrometer. Assignment of ¹H and ¹³C NMR spectra was achieved using 2D methods (COSY, HSQC, HMQC, HMBC). Chemical shifts are given in ppm and were assigned in relation to the solvent signals on the δ-scale^[S1] or to tetramethylsilane (0 ppm) as internal standard. Coupling constants *J* are given in Hertz (Hz). Multiplicities were specified as

follows: s (singlet), d (doublet), dd (double doublet), t (triplet), q (quartet), m (multiplet). For assignment of resonance signals to the appropriate nuclei the following abbreviations were used: Gal (galactose), GlcNAc (*N*-acetylglucosamine), Sia (*N*-acetyl neuraminic acid, sialic acid). Reactions were monitored by TLC using glass plates coated with silica gel 60 F₂₅₄ (Merck) and visualized by using UV light and/or by charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10% H₂SO₄). Column chromatography was performed using automated systems (RediSep Companion or RediSep Rf) from Teledyne Isco with normal phase RediSep columns from the same manufacturer or reversed-phase columns containing LiChroprep RP-18 (40-63 μ m) from Merck KGaA, Darmstadt, Germany. Size exclusion chromatography was performed with Bio-Gel[®] P-2 Gel (45-90 mm) from Bio-Rad. Solvents were purchased from Sigma-Aldrich or Acros. Solvents were dried prior to use where indicated. Tetrahydrofuran (THF) was dried by refluxing with sodium/benzophenone and distilled immediately before use. Dichloromethane (CH₂Cl₂) was dried by filtration over Al₂O₃ (Fluka, type 5016 A basic). Methanol (MeOH) was dried by distillation from sodium methoxide, dimethyl formamide (DMF) by distillation from calcium hydride. Optical rotations were measured using a Perkin-Elmer Polarimeter 341. Electron spray ionization mass spectra (ESI-MS) were obtained on a Waters micromass ZQ. Compound purity was determined on an Agilent 1100 HPLC; detector: ELS, Waters 2420; column: Waters Atlantis dC18, 3 μ m, 4.6 x 75 mm; eluents: A: water + 0.1% TFA; B: 90% acetonitrile + 10% water + 0.1% TFA; linear gradient: 0-1 min 5% B; 1-20 min 5 to 70% B; flow: 0.5 mL/min.

Abbreviations: BzCl, benzoyl chloride; DMAP, 4-dimethylaminopyridine; DMTST, dimethyl (methylthio)sulfonium triflate; MS, molecular sieves; Gal, galactose; GlcNAc, *N*-acetylglucosamine; Sia, *N*-acetyl neuraminic acid/sialic acid; TMSE, 2-(trimethylsilyl)ethyl.

2-(Trimethylsilyl)ethyl 2-acetamido-3-*O*-benzoyl-4,6-*O*-benzylidene-2-deoxy- β -D-glucopyranoside (2). Benzoyl chloride (0.060 mL, 72 mg, 0.52 mmol) and DMAP (cat.) were added to a solution of **1**^[S2] (50 mg, 0.12 mmol) in anhydrous pyridine (0.5 mL) and stirred at r.t. under argon for 4 h. The reaction was quenched with H₂O (20 mL) and the mixture was extracted with CH₂Cl₂ (3 x 15 mL). The combined organic layers were dried (Na₂SO₄) and concentrated. Column chromatography of the residue on silica (petroleum ether/EtOAc) gave **2** (25 mg, 40%) as white solid.

$[\alpha]_D^{20}$ -73.9 (*c* 1.68, CHCl₃); ¹H NMR (500.1 MHz, CDCl₃): δ = 8.11-7.27 (m, 10H, 2 C₆H₅), 6.52 (d, *J* = 9.2 Hz, 1H, NHCOCH₃), 5.67 (t, *J* = 10.0 Hz, 1H, H-3), 5.52 (s, 1H, CHPh), 4.57 (d, *J* = 8.4 Hz, 1H, H-1), 4.35 (m, 1H, H-2), 4.29 (dd, *J* = 4.9, 10.2 Hz, 1H, H-6a), 3.92-3.84 (m, 2H, H-4, OCH₂), 3.80 (t, *J* = 10.2 Hz, H-6b), 3.68 (dt, *J* = 4.9, 10.2 Hz, H-5), 3.41 (m, 1H, OCH₂), 1.86 (s, 3H, NHCOCH₃), 1.00-0.84 (m, 2H, CH₂Si), 0.00 (s, 9H, Si(CH₃)₃); ¹³C NMR (125.8 MHz, CDCl₃): δ = 170.3 (NHCOCH₃), 167.3 (COC₆H₅), 137.2, 133.6, 130.0, 129.4, 129.0, 128.6, 128.2, 126.1 (12C, 2 C₆H₅), 101.6 (C-1), 101.3 (CHPh), 79.2 (C-4), 73.3 (C-3), 68.8 (C-6), 67.6 (OCH₂), 66.2 (C-5), 54.5 (C-2), 23.4 (NHCOCH₃), 18.1 (CH₂Si), -1.3 (3C, Si(CH₃)₃); MS (ESI): *m/z*: Calcd for C₂₇H₃₆NNaO₇Si [M+Na]⁺: 536.2, found: 536.2.

2-(Trimethylsilyl)ethyl 2-acetamido-3-*O*-benzoyl-6-*O*-benzyl-2-deoxy- β -D-glucopyranoside (3). To a solution of **2** (300 mg, 0.58 mmol) in anhydrous THF (12 mL) were added $\text{Me}_3\text{N}\cdot\text{BH}_3$ (170 mg, 2.3 mmol) and anhydrous AlCl_3 (470 mg, 3.5 mmol) with stirring under argon at r.t. After all reagents were dissolved, H_2O (0.01 mL) was added and the mixture was stirred for 5 h.^[S3] The reaction was quenched by addition of H_2O (6 mL) and 1 N aq HCl (6 mL) and extracted with EtOAc (3 x 25 mL). The organic layers were washed with brine, dried over Na_2SO_4 , and concentrated. Column chromatography of the residue on silica gel (petroleum ether/EtOAc, 7:3 to 1:1) gave **3** (238 mg, 0.46 mmol, 79%) as white foam.

^1H NMR (500.1 MHz, CDCl_3): δ = 8.06-7.26 (m, 10H, 2 C_6H_5), 5.81 (d, J = 8.0 Hz, 1H, NHCOCH_3), 5.36 (dd, J = 9.6, 10.6 Hz, 1H, H-3), 4.63-4.61 (m, 2H, H-1, CH_2Ph), 4.58 (B of AB, J = 11.9 Hz, 1H, CH_2Ph), 4.08 (m, 1H, H-2), 3.98 (m, 1H, OCH_2), 3.87 (t, J = 9.6 Hz, 1H, H-4), 3.82 (m, 2H, H-6a, H-6b), 3.65 (dt, J = 4.8, 9.6 Hz, 1H, H-5), 3.56 (m, 1H, OCH_2), 1.85 (s, 3H, NHCOCH_3), 1.05-0.85 (m, 2H, CH_2Si), 0.00 (s, 3H, $\text{Si}(\text{CH}_3)_3$); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 170.4 (NHCOCH_3), 167.7 (COC_6H_5), 137.8, 133.6, 130.1, 129.4, 128.6, 128.0, 127.9 (12C, 2 C_6H_5), 100.6 (C-1), 77.2 (C-3), 76.4 (C-5), 74.3 (CH_2Ph), 73.9 (C-4), 71.3 (C-6), 70.6 (OCH_2), 54.3 (C-2), 23.5 (NHCOCH_3), 18.1 (CH_2Si), -1.3 (3C, $\text{Si}(\text{CH}_3)_3$); MS (ESI): m/z : Calcd for $\text{C}_{27}\text{H}_{37}\text{NNaO}_7\text{Si}$ [$\text{M}+\text{Na}$] $^+$: 538.2, found: 538.2.

2-(Trimethylsilyl)ethyl (benzyl 5-acetamido-4,7,8,9-tetra-*O*-acetyl-3,5-dideoxy-D-glycero- α -D-galacto-2-nonulopyranosynate)-(2 \rightarrow 3)-2,4-di-*O*-acetyl-6-*O*-benzoyl- β -D-galactopyranosyl-(1 \rightarrow 4)-2-acetamido-3-*O*-benzoyl-6-*O*-benzyl-2-deoxy- β -D-glucopyranoside (5). Compounds **3** (200 mg, 0.39 mmol) and **4**^[S4] (190 mg, 0.20 mmol) were dissolved in anhydrous CH_2Cl_2 (5.0 mL) under argon. Activated molecular sieves 4 Å (0.5 g) were added and the mixture was stirred at r.t. for 4 h. Subsequently, a suspension of DMTST (130 mg, 0.50 mmol) and activated molecular sieves 4 Å (0.1 g) in anhydrous CH_2Cl_2 , that had been stirred under argon at r.t. for 4 h, was added. After 5 d, the reaction mixture was diluted with CH_2Cl_2 (10 mL), filtered through a pad of celite, and washed with satd. aq. NaHCO_3 (20 mL) and brine (20 mL). The aqueous layers were extracted with CH_2Cl_2 (3 x 20 mL) and the combined organic layers were dried over Na_2SO_4 and concentrated. Column chromatography of the residue on silica (petroleum ether/EtOAc/MeOH, 8:5:0.5 to 8:5:0.9) gave **5** (109 mg, 0.08 mmol, 39%) as white solid.

^1H NMR (500.1 MHz, CDCl_3): δ = 8.13-7.21 (m, 20H, 4 C_6H_5), 5.68 (d, J = 9.4 Hz, 1H, GlcNAc-NHCOCH_3), 5.55 (ddd, J = 2.7, 5.9, 8.8 Hz, 1H, Sia-H8), 5.39 (dd, J = 8.3, 9.5 Hz, 1H, GlcNAc-H3), 5.30 (dd, J = 2.7, 8.8 Hz, 1H, Sia-H7), 5.25 (A of AB, J = 12.0 Hz, 1H, CH_2Ph), 4.99-4.94 (m, 2H, Gal-H2, Gal-H4), 4.91-4.78 (m, 4H, Gal-H1, Sia-H4, Sia- NHCOCH_3 , CH_2Ph), 4.64-4.59 (m, 2H, CH_2Ph), 4.58 (dd, J = 3.5, 10.0 Hz, 1H, Gal-H3), 4.51 (d, J = 7.5 Hz, 1H, GlcNAc-H1), 4.32 (dd, J = 2.7, 12.4 Hz, 1H, Sia-H9a), 4.23-4.11 (m, 2H, GlcNAc-H2 , GlcNAc-H4), 4.05-3.94 (m, 3H, Sia-H5, Sia-H9b, OCH_2), 3.88-3.65 (m, 6H, Gal-H5, Gal-H6a, Gal-H6b, GlcNAc-H5 , GlcNAc-H6a , GlcNAc-H6b), 3.56 (m, 1H, OCH_2), 3.45 (dd, J = 2.7, 10.7 Hz, Sia-H6), 2.56 (dd, J = 4.6, 12.7 Hz, Sia-H3_{eq}), 2.13, 2.10, 2.03, 2.02, 1.97, 1.93, 1.87, 1.82 (8 s, 24H, 8 COCH_3), 1.65 (m, 1H, Sia-H3_{ax}), 1.02-0.90 (m, 2H, CH_2Si), 0.01 (s, 9H, $\text{Si}(\text{CH}_3)_3$); ^{13}C NMR (125.8 MHz, CDCl_3): δ = 170.7, 170.4, 170.2, 170.1, 170.0, 169.7 (8C, 8 COCH_3), 167.5 (Sia-C1), 166.4, 165.4 (2 COC_6H_5), 138.5, 134.8, 133.4, 133.3, 130.1, 129.9, 129.8, 128.9, 128.8, 128.7, 128.6, 128.5, 128.4, 127.6 (24C, Ar-C), 100.9 (Gal-C1), 100.7 (GlcNAc-C1), 96.9 (Sia-C2), 75.4 (GlcNAc-C4), 75.0 (GlcNAc-C5), 73.7, 73.6 (GlcNAc-C3 , CH_2Ph), 72.2 (Sia-C6), 71.6 (Gal-C3), 70.7, 70.6 (Gal-C2, Gal-

C5), 69.4 (Sia-C4), 68.8 (Gal-C6), 68.5 (CH₂Ph), 67.8 (Sia-C8), 67.2, 67.1, 66.8 (Sia-C7, Gal-C4, OCH₂), 62.6 (Sia-C9), 61.0 (GlcNAc-C6), 53.6 (GlcNAc-C2), 49.0 (Sia-C5), 37.5 (Sia-C3), 23.4, 23.3, 21.5, 21.1, 20.9, 20.9, 20.9, 20.7 (8 COCH₃), 18.2 (CH₂Si), -1.3 (3C, Si(CH₃)₃); MS (ESI): *m/z*: Calcd for C₇₀H₈₆N₂NaO₂₇Si⁺ [M+Na]⁺: 1437.5, found: 1437.8.

2-(Trimethylsilyl)ethyl (sodium 5-acetamido-3,5-dideoxy-D-glycero-α-D-galacto-2-nonulopyranosynate)-(2→3)-β-D-galactopyranosyl-(1→4)-2-acetamido-2-deoxy-β-D-glucopyranoside (6). Pd(OH)₂/C (15 mg, 10% Pd) was added to a solution of **5** (100 mg, 0.07 mmol) in dioxane/H₂O (4:1, 5.0 mL) and the resulting mixture was hydrogenated (1 bar H₂) at r.t. After 2 d, the reaction mixture was filtered, concentrated, and the residue was redissolved in MeOH (1.0 mL) and aq. 1.0 M NaOH (5.0 mL). After stirring for 20 h at r.t., the solution was neutralized with aq. 1.0 M HCl and concentrated. Size exclusion chromatography of the residue, followed by reversed phase chromatography (RP-18, H₂O/MeOH) gave **6** (43 mg, 0.54 mmol, 77%) as white foam.

[α]_D²⁰ -14.1 (*c* 0.99, MeOH); ¹H NMR (500.1 MHz, D₂O/CD₃OD): δ = 4.55 (d, *J* = 7.9 Hz, 2H, Gal-H1, GlcNAc-H1), 4.11 (dd, *J* = 3.1, 9.9 Hz, 1H, Gal-H3), 4.04 (td, *J* = 4.9, 10.4 Hz, 1H, OCH₂), 3.99 (dd, *J* = 2.1, 12.3 Hz, GlcNAc-H6a), 3.95 (d, *J* = 3.1 Hz, 1H, Gal-H4), 3.92-3.54 (m, 16H, OCH₂, Gal-H2, Gal-H5, Gal-H6a, Gal-H6b, GlcNAc-H2, GlcNAc-H3, GlcNAc-H4, GlcNAc-H6b, Sia-H4, Sia-H5, Sia-H6, Sia-H7, Sia-H8, Sia-H9a, Sia-H9b), 2.76 (dd, *J* = 4.6, 12.1 Hz, 1H, Sia-H3_{eq}), 2.03, 2.02 (2 s, 6H, 2 COCH₃), 1.80 (t, *J* = 12.1 Hz, 1H, Sia-H3_{ax}), 1.05-0.80 (m, 2H, CH₂Si), 0.00 (s, 9H, Si(CH₃)₃); ¹³C NMR (125.8 MHz, D₂O/CD₃OD): δ = 176.0 (Sia-C1), 175.3, 174.8 (2 COCH₃), 103.5 (Sia-C2), 101.3 (GlcNAc-C1), 100.8 (Gal-C1), 79.3 (GlcNAc-C3), 76.5 (Gal-C3), 76.1 (GlcNAc-C5), 75.8 (Gal-C5), 73.9, 73.7, 72.7, 70.4, 69.3, 69.3, 69.1 (Gal-C2, Sia-C4, Sia-C6, Sia-C7, Sia-C8, GlcNAc-C4, OCH₂), 68.5 (Gal-C4), 63.6 (Sia-C9), 62.0 (Gal-C6), 61.0 (GlcNAc-C6), 56.0 (GlcNAc-C2), 52.7 (Sia-C5), 40.6 (Sia-C3), 23.2, 23.0 (2 COCH₃), 18.1 (CH₂Si), -1.5 (3C, Si(CH₃)₃); HR-MS (ESI): *m/z*: Calcd for C₃₀H₅₄N₂NaO₁₉Si [M+H]⁺: 797.2982, found: 797.2983; HPLC purity: > 99.5%.

2. LC-MS/MS analysis of E-selectin

To confirm the mutation at the amino acid level, purified wt and S128R E-selectin were submitted to LC-MS/MS analysis. Prior to the analysis, the proteins were deglycosylated with PNGaseF under denaturing conditions according to the manufacturers protocol (New England BioLabs).

The protein band was reduced prior to digestion with 10 mM dithiothreitol for 2 h at 37 °C and alkylated with iodoacetamide at 50 mM final concentration for 15 min at r.t. in the dark. The gel piece was digested with 125 ng trypsin (sequencing grade) for 18 h at 37 °C. The peptides in the supernatant were collected and the gel piece was extracted with aqueous solution of 0.1% acetic acid / 50% acetonitrile and the extract was pooled with the tryptic peptides. The digest was dried in a SpeedVac and redissolved in 0.1% acetic acid / 2% acetonitrile.

The trypsin digests were analyzed by capillary liquid chromatography tandem MS (LC-MS/MS) using a set up of a trapping 300SB C-18 column (0.3 x 50 mm) and a separating column (0.1 mm x 10 cm) that had been packed with Magic 300 Å C-18 reversed-phase

material (5 μm particle size). The columns were connected on line to an Orbitrap FT hybrid instrument. A linear gradient from 2 to 80% solvent B (0.1% acetic acid and 80% acetonitrile in water) in solvent A (0.1% acetic acid and 2% acetonitrile in water) was delivered with a Rheos 2200 pump at a flow rate of 100 $\mu\text{L}/\text{min}$ during 85 min. A pre-column split was used to reduce the flow to approximately 100 nL/min. 10 μL of peptide digest were injected with an autosampler thermostatted to 4 $^{\circ}\text{C}$. The eluted peptides were ionized at 1.7 kV. The mass spectrometer was operated in a data-dependent fashion. The precursor scan was done in the Orbitrap set to 60,000 resolution, while the fragment ions were mass analyzed in the LTQ instrument. A top five method was run so that the five most intense precursors were selected for fragmentation. The MS/MS spectra were then searched using TurboSequest software.^[S5] The data were searched against the protein itself with Bioworks version 3.3.1. SP1 by setting the precursor ion tolerance to 10 ppm, while the fragment ion tolerance was set to 0.5 Da. Cleavage rules were set to fully enzymatic-cleaves at both ends, allowing 2 missed cleavages. Post filtering was set to the following parameters: ΔCN , 0.1; X_{corr} versus charge state was 1.50 (1+), 2.00 (2+), 2.50 (3+); peptide probability, 0.01; protein probability 0.01

Fragment with the sequence K.KLALCYTAACTNTSCR.G or K.LALCYTAACTNTSCR.G were found which correspond to the S128R mutation.

References

- [S1] Gottlieb HE, Kotlyar V, Nudelman A. 1997. NMR chemical shifts of common laboratory solvents as trace impurities. *J. Org. Chem.*, 62:7512-7515.
- [S2] Takaku H, Sato J, Ishida H, Inazu T, Ishida H, Kiso M. 2006. A chemical synthesis of UDP-LacNAc and its regioisomer for finding 'oligosaccharide transferases'. *Glycoconj. J.*, 23:565-573.
- [S3] Sherman AA, Mironov YV, Yudina ON, Nifantiev NE. 2003. The presence of water improves reductive openings of benzylidene acetals with trimethylaminoborane and aluminium chloride. *Carbohydr. Res.*, 338:697-703.
- [S4] Bhunia A, Schwardt O, Gathje H, Gao GP, Kelm S, Benie AJ, Hricovini M, Peters T, Ernst B. 2008. Consistent Bioactive Conformation of the Neu5Ac $\alpha(2 \rightarrow 3)\text{Gal}$ Epitope Upon Lectin Binding. *Chembiochem*, 9:2941-2945.
- [S5] Gatlin CL, Eng JK, Cross ST, Detter JC, Yates JR. 2000. Automated identification of amino acid sequence variations in proteins by HPLC/microspray tandem mass spectrometry. *Anal. Chem.*, 72:757-763.

Chapter 2.6. – Manuscript 3

***In vitro* Comparison of Murine and Human E-selectin Binding to Natural Ligands and Antagonists**

Manuscript in preparation. Written as short communication.

Contributions

- | | |
|------------------|---|
| R.C. Preston | <ul style="list-style-type: none">• Project planning and supervision.• Protein expression and purification.• Assay development.• Manuscript preparation. |
| S. Röthlisberger | <ul style="list-style-type: none">• Protein expression and purification.• Assay performance. |
| Ganpan Gao | <ul style="list-style-type: none">• Synthesis of <i>N</i>-glycolyl-sLe^x. |

In vitro* Comparison of Murine and Human E-selectin Binding to Natural Ligands and Antagonists*

Roland C. Preston¹, Silvan Röthlisberger¹, Ganpan Gao¹, John L. Magnani², Beat Ernst^{1*}

¹ Institute of Molecular Pharmacy, Pharmacenter, University of Basel, 4056 Basel, Switzerland

² Glycomimetics Inc. Gaithersburg, MD 20879, USA.

[*]R. C. Preston, S. Röthlisberger, Dr. G. Gao, Prof. Dr. B. Ernst

Institute of Molecular Pharmacy, University of Basel

Klingelbergstrasse 50 CH-4056 Basel (Switzerland)

E-mail : beat.ernst@unibas.ch

Dr. J. L. Magnani

GlycoMimetics, Inc., Gaithersburg, Maryland 20879 (USA)

Abbreviations: *CMAH*, gene encoding cytidine monophosphate-*N*-acetylneuraminic acid hydroxylase; D-Gal, D-galactose; D-GlcNAc, *N*-Acetylglucosamine; D-Neu5Ac, *N*-Acetylneuraminic acid; D-Neu5Gc, *N*-Glycolylneuraminic acid; EGF, epidermal growthfactor-like domain, IgG(Fc), fragment crystallizable of immunoglobulin G; Lec, lectin domain; L-Fuc, L-fucose; PAA, polyacrylamide; SCR, short consensus repeats; sLe^{x/a}, sialyl Lewis^{x/a}.

[**] The authors gratefully acknowledge the financial support of the Swiss National Science Foundation (200020_129935/1). We further thank Prof. Dr. Vestweber (Max Planck Institute for Molecular Biomedicine, Münster, Germany) for providing murine E-selectin expressing CHO cells.

Keywords Human E-selectin; Murine E-selectin; Selectin Antagonists; Glycomimetics.

Abstract

E-selectin is a human cell-adhesion molecule which is involved in leukocyte recruitment during inflammation. Dysregulation of E-selectin may, however, lead to pathological situations, *e.g.* ischemic stroke, asthma, or psoriasis. Glycomimetic E-selectin antagonists based on the natural ligand sialyl Lewis^x (sLe^x) have been developed to prevent E-selectin mediated leukocyte recruitment. Such antagonists have been developed under guidance of *in vitro* target- and cell-based assays with the human form of E-selectin. For the approval of clinical studies, *in vivo* efficacy needs to be demonstrated. For selectin-related diseases, various mouse models exist, but the affinity of E-selectin antagonists towards murine E-selectin was never investigated *in vitro* to date.

In this study we compared binding affinities of the natural ligand sLe^x and the isoform sLe^a to recombinantly expressed and purified human and murine E-selectin. Furthermore, we compared binding of *N*-glycolyl-sLe^x (**1**), the natural ligand for E-selectin in mice. Finally, we assessed binding affinities of three sLe^x (**2–4**) mimetics to murine E-selectin.

For sLe^x, a three-fold higher affinity was observed for murine E-selectin when compared to the human counterpart, while sLe^a and *N*-glycolyl-sLe^x (**1**) showed comparable affinities. All three sLe^x glycomimetics (**2–4**) were capable of binding to murine E-selectin and exhibited three- to five-fold increased affinity compared to human E-selectin.

We therefore confirm that mouse models are suitable for the evaluation of E-selectin antagonists *in vivo*. However, the effect of these antagonists may be overestimated, since generally higher affinities for murine E-selectin were measured.

Introduction

The family of selectins with its three members E-, P- and L-selectin are transmembrane calcium dependent lectins. They act as vascular cell-adhesion molecules that are responsible for leukocyte trafficking.^[1] Being part of the innate immune system, E- and P-selectin appear on post-capillary venules upon cytokine stimulation upon tissue injury and form the initial contact to leukocytes circulating in the blood stream. This leukocyte tethering is the initial step of an entire chain of events, leading to rolling of leukocytes along endothelial vasculature, firm adhesion, and transmigration into the surrounding tissue where leukocytes counteract the cause of inflammation. To allow tethering and rolling, the selectins interact with carbohydrates that are presented on the cell surface of leukocytes. The presence of the tetrasaccharide sialyl Lewis^x (sLe^x, Neu5Ac(α 2-3)Gal(β 1-4)[Fuc(α 1-3)]GlcNAc) on glycoproteins is an absolute necessity for selectin mediated interaction in humans. Unlike most species, humans are not capable of converting *N*-acetyl-neuraminic acid (Neu5Ac) into *N*-glycolyl-neuraminic acid (Neu5Gc) during sialic acid biosynthesis. This is a result of a deletion of the *CMAH* gene in humans that encodes the cytidine monophosphate-*N*-acetylneuraminic acid hydroxylase.^[2] Therefore, *N*-glycolyl-sLe^x (**1**) is the natural ligand for E-selectin in most species, also for mice.

While the presentation of E-selectin on vascular endothelial cells is strictly regulated on the transcription level in healthy tissue, dysregulation of immune processes can lead to locally excessive leukocyte accumulation. The massive influx of leukocytes and their release of cytotoxic agents can subsequently lead to leukocyte-inflicted tissue damage. Such an inadequate immune response over a prolonged time period can lead to various diseases, *e.g.* atherosclerosis, reperfusion injury, stroke and psoriasis. By targeting the initial adhesion step with selectin antagonists, the emergence of selectin-mediated diseases can be prevented. Efforts in drug discovery to target selectins have lead to therapeutic antibodies,^[3] recombinant glycoprotein fragments,^[4] or small-molecule antagonists, *e.g.* glycomimetics that are based on sLe^x as lead structure.^[5] For therapeutic use of small-molecule antagonists, their safety and efficacy are required to be assessed in pre-clinical trials, which involves proof-of-concept studies in animal models. For glycomimetic E-selectin antagonists, leukocyte rolling and the inhibition thereof is investigated by intravital microscopy in C57BL/6 mice.^[5a] This mouse model requires the induction of selectin presentation which is achieved by injection of tumor necrosis factor- α .^[6] For vascular occlusion models, sickle cell disease mice are generated by bone marrow transplants from sickle cell mice into irradiated

C57BL/6 mice.^[5b] *In vitro* analysis of glycomimetic E-selectin antagonists are performed with a set of techniques, such as flow chamber assays,^[7] target-based polymer-binding assays,^[5e, 7-8] surface plasmon resonance,^[5d] and isothermal titration calorimetry.^[9] In these assays, human E-selectin expressing cells (human umbilical vein cells) or recombinantly expressed and purified human E-selectin is employed. Therefore, a correlation of affinity data originating from *in vitro* studies using human E-selectin and *in vivo* mouse models with murine E-selectin as target might not be straight forward.

In this study we assessed the affinity of glycomimetic E-selectin antagonist binding to murine E-selectin in comparison to human E-selectin *in vitro*. While the terminal residue in sLe^x in mice is either Neu5Ac or Neu5Gc, it is exclusively Neu5Ac in humans. It is therefore conceivable that human E-selectin has adapted to the *N*-acetyl form of sLe^x and that modifications of sLe^x based mimics, especially of the Neu5Ac portion, might exhibit differential impact on murine E-selectin in comparison to human E-selectin. The distinction of the two sialic acids has previously been studied on Siglec-1, a sialic acid binding member of the immunoglobulin family, where the human protein strongly prefers Neu5Ac over Neu5Gc.^[10]

```

Human E-selectin   1 WSYNTSTEAMTYDEASAYCQQRYTHLVAIQNKEEIEYLNSILSYSPSYWIGIRKVNNVW
Murine E-selectin  1 WYYNASSELMTYDEASAYCQRDYTHLVAIQNKEEINYLSNLKHSPSYWIGIRKVNNVW
                  * * * * *
Human E-selectin   61 VWVGTQKPLTEEAKNWAPGEPNNRQKDEDCVEIYIKREKDVGMWNDERCSKKKLALCYTA
Human E-selectin   61 IWVGTGKPLTEEAQNWAPGEPNNKQRNEDCVEIYIQRTKDSGMWNDERCNKKKLALCYTA
                  **** *
Human E-selectin  121 ACTNTSCSGHGECVETINYTCKCDPGFSGLKCEQIV
Human E-selectin  121 SCTNASCSGHGECIETINYTCKCHPGFLGPNCEQAV
                  *** *

```

Figure 1. Sequence alignment of the lectin domain (residues 1–120) and the EGF-like domain (residues 121–156) of human and murine E-selectin. Asparagin residues marked green are potential *N*-linked glycosylation sites (as determined by NetNGlyc v1.0). Residues marked in red are involved in ligand binding as observed in the crystal structure of human E-selectin in complex with sLe^x (PDB code 1G1T).^[11] The lectin domains share 83% sequence identity, while the EGF-like domains share 77% identity.

Human and murine E-selectin share an identical overall structure, with an N-terminal lectin (Lec) domain containing the carbohydrate binding site, an epidermal growth factor-like (EGF) domain, six short consensus repeats (SCR), a transmembrane region, and a cytoplasmic domain. The sequence homology across all domains is 73%. The Lec and EGF domains, which together provide the minimal requirement for carbohydrate binding, are more conserved with 82% sequence similarity (Figure 1). The amino acid residues in direct contact

with sLe^x or the Ca²⁺ ion, as determined by X-ray crystallography of the human E-selectin/sLe^x complex,^[11] are fully conserved in murine E-selectin. However, two residues in close proximity to Neu5Ac are not conserved. The Tyr44 and Glu98 in human E-selectin are exchanged by a His44 and Thr98 in murine E-selectin (Figure 2).

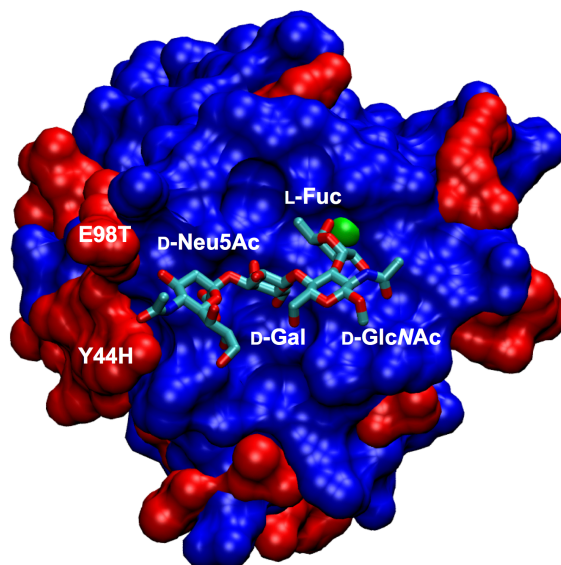


Figure 2. Surface representation of human E-selectin in complex with sLe^x (PDB code 1G1T).^[11] Residues colored in blue are conserved in human and murine E-selectin, while residues colored in red mark non-conserved residues. The Ca²⁺ ion is depicted as green sphere. While most of the binding site is well conserved, two residues near the D-Neu5Ac moiety differ.

Results and Discussion

To compare human and murine E-selectin *in vitro*, we expressed the Lec (L), EGF (E) and the first two SCRs (C2) of both types of E-selectin as IgG_{FC} constructs in Chinese hamster ovary cells. Both proteins were purified to homogeneity (Figure 3A). While the human E-selectin_{LEC2}-IgG_{FC} construct shows an apparent molecular weight of 95 kDa, murine E-selectin_{LEC2}-IgG_{FC} exhibits a lower apparent molecular weight of 80 kDa. This suggests different glycosylation patterns of human and murine E-selectin. Indeed, while human E-selectin is predicted to possess seven potential *N*-linked glycosylation sites in the first four domains (predicted by NetNGlyc v1.0, threshold = 0.5), murine E-selectin has only five potential sites. The position of four *N*-linked glycosylation sites are conserved in the first four domains (Asn4, Asn124, Asn178, and Asn244).

The affinity of the natural ligands sLe^x and its isoform sLe^a (Neu5Acα2→3Galβ1→3(Fucα1→4)GlcNAc), also an E-selectin ligand,^[12] was compared in a target-based binding assay with peroxidase-conjugated carbohydrate polyacrylamide (sLe^x-PAA/sLe^a-PAA) as shown in Figure 3B.^[8a, 8b] For human E-selectin_{LEC2}-IgG_{FC}, the

EC_{50} value of sLe^x -PAA was 315 ± 63 ng/mL, whereas sLe^a -PAA showed a threefold improved affinity of 112 ± 5 ng/mL. The higher affinity for sLe^a compared to sLe^x for human E-selectin has previously been observed in *in vitro* assays.^[8a, 13] For murine E-selectin_{LEC2}-IgG_{FC}, the EC_{50} value of sLe^x (105 ± 8 ng/mL) was similar to that of sLe^a (96 ± 9 ng/mL). Therefore, sLe^x -PAA exhibited a threefold higher affinity to murine E-selectin than to the human form. For sLe^a -PAA the affinities were similar for both, human and murine E-selectin. Human E-selectin is therefore capable of distinguishing $\alpha 1 \rightarrow 3$ and $\alpha 1 \rightarrow 4$ linkage of L-Fuc, while murine E-selectin binds both isoforms with similar affinities.

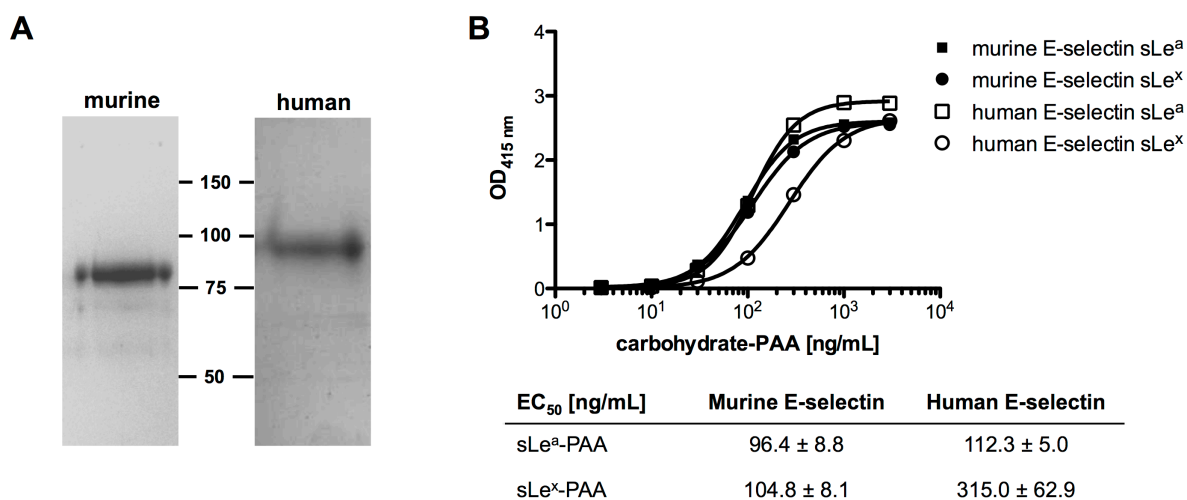


Figure 3. (A) SDS-PAGE analysis of purified human and murine E-selectin_{LEC2}IgG_{FC}. (B) Binding of sLe^x - and sLe^a -polyacrylamide-peroxidase (PAA) to immobilized human and murine E-selectin. Signals were obtained after color reaction with a peroxidase substrate. EC_{50} values are the concentration of carbohydrate-PAA at which half-maximal binding was observed. Errors given are standard deviations from two independent measurements, each run in duplicate wells.

We assessed the affinity of *N*-glycolyl- sLe^x (**1**), which is the natural ligand to E-selectin in mice, with a competitive inhibition assay (Figure 4). Since sLe^a -PAA exhibited a similar affinity for the human and murine form of E-selectin, this ligand was chosen as reporter in the competitive binding assays. Binding of *N*-glycolyl- sLe^x (**1**) to human E-selectin was in the millimolar range ($IC_{50} = 1.7$ mM), similar as observed for *N*-acetyl- sLe^x .^[14] The affinity of *N*-glycolyl- sLe^x (**1**) to murine E-selectin ($IC_{50} = 1.4$ mM) was similar as observed for the human counterpart. Murine E-selectin is therefore not adapted to the Neu5Gc form of sLe^x . Since no interactions of Neu5Ac are observed in the E-selectin/ sLe^x crystal structure except for the carboxylic acid, it is not surprising that the exchange of an acetyl to glycolyl group does not influence binding. It is noteworthy, however, that murine E-selectin/IgG bound

N-acetyl-sLe^x-PAA three-fold better than the human counterpart and that this appears not to be the case for *N*-glycolyl-sLe^x, albeit being in the $\alpha 1 \rightarrow 3$ configuration for L-Fuc.

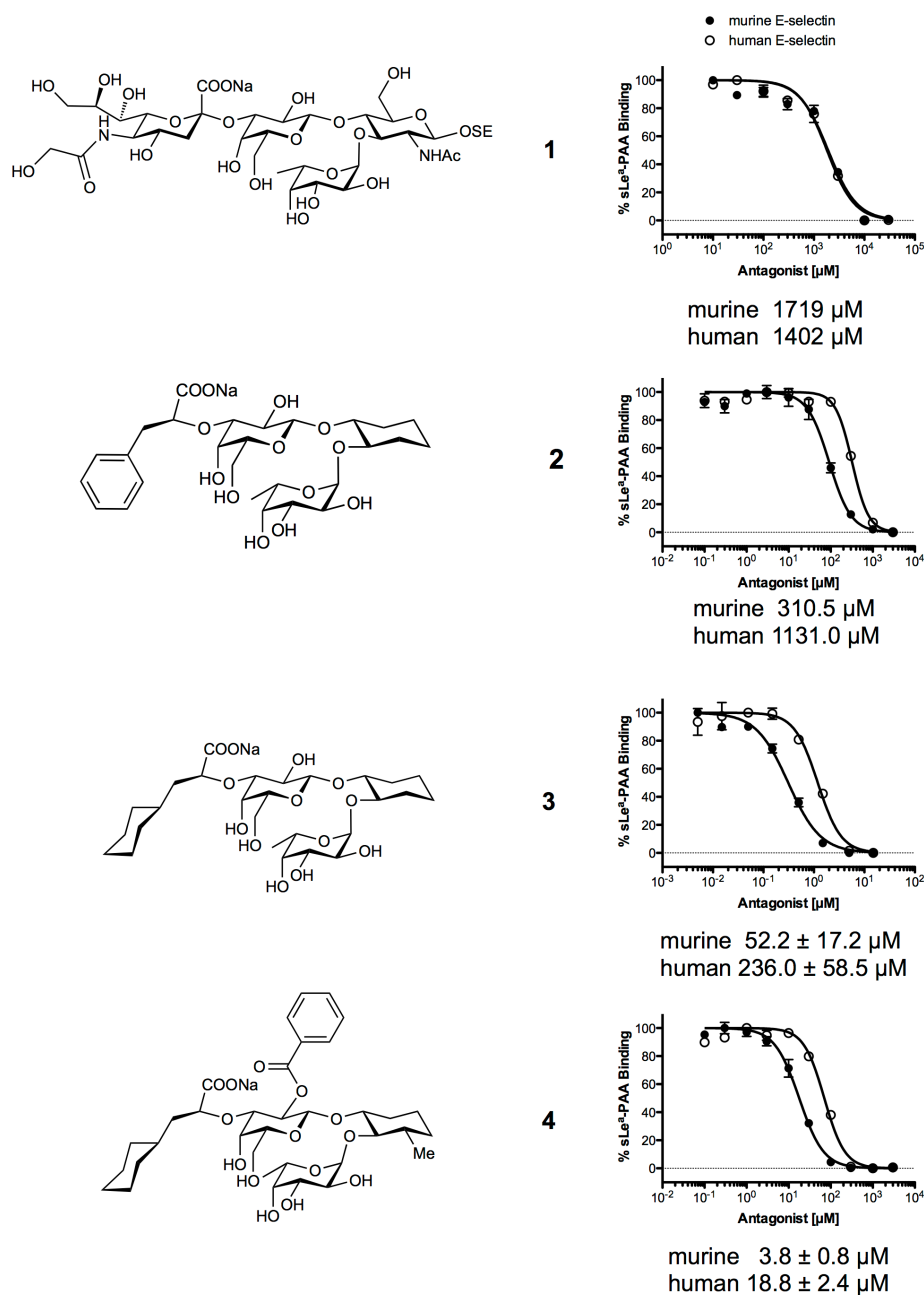


Figure 4. Binding affinities of the natural ligand for murine E-selectin (*N*-glycolyl-sLe^x, **1**) and the three sLe^x mimetics **2-4**. Binding affinities were measured in a target-based competitive binding assay with a dilution series of each compound. Compounds **1** and **2** were measured a single time in duplicate wells. The errors for compound **3** and **4** are standard deviations of two independent measurements.

To evaluate the efficacy of sLe^x derived glycomimetics with murine E-selectin, selected compounds were tested in the target-based inhibition assay. Three glycomimetic compounds that cover a wide range of affinities (20–1500 μM) were tested for their affinity.^[5d, 9] In compound **2**, D-GlcNAc and D-Neu5Ac are replaced by (1*R*,2*R*)-cyclohexane-1,2-diol and

(*S*)-phenyl lactic acid, respectively.^[5a, 15] Compound **2** previously exhibited a 4.5-fold weaker affinity than compound **3** in a target-based inhibition assay.^[8c] In compound **3** (**CGP69669A**), Neu5Ac has been exchanged with (*S*)-cyclohexyl lactic acid, while otherwise being identical to compound **2**. The affinity of compound **3** to human E-selectin_{LEC6}IgG_{Fc} was shown to be 12-fold improved over sLe^x.^[5d] The glycomimetic compound **4** consists of the same replacement for D-Neu5Ac as compound **3**. Furthermore, the D-GlcNAc moiety is replaced by (1*R*,2*R*,3*S*)-3-methylcyclohexane-1,2-diol, and the D-Gal moiety is additionally benzoylated in the 2'-position. Compound **4** was shown to be 15 to 20-fold more potent than compound **3**.^[5d]

For the human E-selectin_{LEC2}IgG_{Fc} construct, similar relative IC₅₀ values were obtained as reported in the literature. Compared to compound **3** (236 ± 58 μM), compound **4** was 13-fold more potent (19 ± 2 μM), while compound **2** was 5-fold weaker (1131 μM). However, the absolute values were higher than previously reported. For example, the IC₅₀ value of compound **3** to E-selectin_{LEC6}IgG_{Fc} is reported to be 80 μM, whereas a value of 236 ± 58 μM was measured with the E-selectin_{LEC2}IgG_{Fc} construct. The difference in affinities is a consequence of the reduced number of SCR domains, which has been shown to influence E-selectin activity, but not its selectivity.^[16] The glycomimetic compounds were also able of inhibiting sLe^a-PAA binding to murine E-selectin. A similar ranking of the compounds was observed. In relation to compound **3** (52 ± 17 μM), compound **4** was 14-fold more potent (3.8 ± 0.8 μM), and compound **2** was 6-fold weaker (310 μM). In direct comparison, the affinities of all glycomimetics were 3.5–5-fold improved for murine E-selectin than for human E-selectin. However, the differences in affinities was consistent with the comparison of the EC₅₀ values of the sLe^x-PAA, where murine E-selectin exhibited 3-fold increased binding potency. In that regard, the augmented affinity to murine E-selectin is not surprising since the tested glycomimetics are based on sLe^x and therefore all possess α1→3 linked L-Fuc moieties.

Conclusion and Summary

In summary, we show that sLe^x glycomimetics designed against human E-selectin are also capable of binding to murine E-selectin. This piece of information is important since *in vivo* efficacy studies are routinely performed in mice. Based on the herein presented data, mouse studies with E-selectin antagonists are suitable for proof-of-concept studies. However, since the affinity of antagonists was approximately three-fold higher for murine E-selectin compared to human E-selectin, this must be taken into consideration when correlating dosage for human application.

Material and Methods

Protein expression and purification

CHO-K1 cells expressing human E-selectin expressing Lec-EGF-SCR2 domains fused to the Fc-part of human IgG2 was generated as previously reported.^[17] Protein was purified with a protein A-Sepharose[®] column and by an anti-human E-selectin antibody (7A9) as previously described.^[9] CHO-K1 cells expressing murine E-selectin Lec-EGF-SCR2 domains fused to the Fc-part of murine IgG1 were a generous gift from Prof. Dr. Vestweber (Max Planck Institute for Molecular Biomedicine, Münster, Germany). The cells were grown as monolayers at 37 °C and 5% CO₂ in Ham's Nutrient Mixture F-12 medium supplemented with 2 mM L-glutamine, 10% (v/v) fetal calf serum, 100 U/mL penicillin, and 100 µg/mL streptomycin. Once confluent, culture medium was collected weekly. Murine E-selectin was initially purified with a protein A-sepharose[®] according to the human counterpart. The protein was subsequently purified to homogeneity by anion exchange chromatography with a Uno[™] Q6 column (Bio-Rad, Reinach BL, Switzerland) and buffers A (20 mM Tris/HCl, pH 8.8) and B (20 mM Tris/HCl, pH 8.8, 1 M NaCl). Protein was concentrated by ultrafiltration (50 kDa cut-off, Millipore AG, Zug, Switzerland).

Target-based binding assays

The target-based binding assays were performed similar as previously reported.^[17] Briefly, 20 µL biotinylated sLe^x- and sLe^a-PAA (100 µg/mL, GlycoTech, Gaithersburg, MD) was incubated together with 80 µL streptavidinidase (200 U/mL, Roche Applied Science, Rotkreuz, Switzerland), 10 µL fetal calf serum and 80 µL assay buffer (20 mM HEPES/NaOH pH 7.4, 150 mM NaCl, and 1 mM CaCl₂) at 37 °C for 2 h. Human and murine E-selectin/IgG diluted to 3 µg/mL in assay buffer was passively coated on Immunoplate II MaxiSorp[™] 96-well microtiter plates (100 µL/well, Nunc, Roskilde, Denmark) overnight at 4 °C. Wells were washed with assay buffer (150 µL/well) and blocked with 2% BSA in assay buffer (200 µL/well). For EC₅₀ assays, serial dilutions (3 ng/mL–3 µg/mL) of the carbohydrate-PAA was incubated for 3 h at 25 °C while shaking (300 rpm). For competitive inhibition assays, serial dilutions of the compounds (**1–4**) were prepared with 10% DMSO (v/v) and diluted 1:1 with sLe^a-PAA (200 ng/mL) prior to incubating on the protein-coated wells. After washing and removal of excess carbohydrate-PAA, 2,2'-azino-di-(3-ethylbenzthiazoline)-6-sulfonic acid (100 µL/well, Invitrogen, Lucerne, Switzerland) was added for colorimetric reaction and stopped by addition of 2% oxalic acid (w/v, 100 µL/well)

after 1–2 minutes. The optical density was measured at 415 nm. EC₅₀ and IC₅₀ values were determined with the Prism software (GraphPad Software Inc., La Jolla, CA). A four-parameter fit for minimum, maximum, logIC₅₀/logEC₅₀ and slope was used.

Synthesis of compounds

Compounds **1-4** were synthesized as described previously.^[5d, 8c, 18]

References

- [1] K. Ley, G. S. Kansas, *Nat. Rev. Immunol.* **2004**, *4*, 325-335.
- [2] a) A. Irie, S. Koyama, Y. Kozutsumi, T. Kawasaki, A. Suzuki, *J. Biol. Chem.* **1998**, *273*, 15866-15871; b) A. Varki, *Yearb Phys Anthropol* **2001**, *44*, 54-69.
- [3] A. Seekamp, M. van Griensven, E. Dhont, M. Diefenbeck, I. Demeyer, G. Vundelinckx, N. Haas, U. Schaechinger, L. Wolowickwa, S. Rammelt, J. Stroobants, I. Marzi, A. M. Brambrink, P. Dziurdzik, J. Gasiorowski, H. Redl, M. Beckert, J. Khan-Boluki, *Crit. Care Med.* **2004**, *32*, 2021-2028.
- [4] K. Wang, X. R. Zhou, Z. M. Zhou, K. Tarakji, J. X. Qin, M. Sitges, T. Shiota, F. Forudi, R. G. Schaub, A. Kumar, M. S. Penn, E. J. Topol, A. M. Lincoff, *Thromb. Haemost.* **2002**, *88*, 149-154.
- [5] a) K. E. Norman, G. P. Anderson, H. C. Kolb, K. Ley, B. Ernst, *Blood* **1998**, *91*, 475-483; b) J. S. Chang, J. T. Patton, A. Sarkar, B. Ernst, J. L. Magnani, P. S. Frenette, *Blood* **2010**, *116*, 1779-1786; c) J. Patton, A. Sarkar, B. Wagner, J. S. Chang, P. Frenette, H. Thackray, B. Ernst, J. L. Magnani, *Glycobiology* **2009**, *19*, 1332-1332; d) D. Schwizer, J. T. Patton, B. Cutting, M. Smiesko, B. Wagner, A. Kato, C. Weckerle, F. P. Binder, S. Rabbani, O. Schwardt, J. L. Magnani, B. Ernst, *Chemistry* **2012**, *18*, 1342-1351; e) A. Titz, J. Patton, M. Smiesko, Z. Radic, O. Schwardt, J. L. Magnani, B. Ernst, *Bioorg. Med. Chem.* **2010**, *18*, 19-27.
- [6] K. Ley, D. C. Bullard, M. L. Arbones, R. Bosse, D. Vestweber, T. F. Tedder, A. L. Beaudet, *J. Exp. Med.* **1995**, *181*, 669-675.
- [7] G. Thoma, J. L. Magnani, J. T. Patton, B. Ernst, W. Jahnke, *Angew. Chem. Int. Ed.* **2001**, *40*, 1941-1945.
- [8] a) G. Weitz-Schmidt, D. Stokmaier, G. Scheel, N. E. Nifantev, A. B. Tuzikov, N. V. Bovin, *Anal. Biochem.* **1996**, *238*, 184-190; b) G. Thoma, J. L. Magnani, R. Ohrlein, B. Ernst, F. Schwarzenbach, R. O. Duthaler, *J. Am. Chem. Soc.* **1997**, *119*, 7414-7415; c) R. Banteli, B. Ernst, *Bioorg. Med. Chem. Lett.* **2001**, *11*, 459-462.
- [9] F. P. Binder, K. Lemme, R. C. Preston, B. Ernst, *Angew. Chem. Int. Ed. Engl.* **2012**, *51*, 7327-7331.
- [10] E. C. M. Brinkman-Van der Linden, E. R. Sjoberg, L. R. Juneja, P. R. Crocker, N. Varki, A. Varki, *J. Biol. Chem.* **2000**, *275*, 8633-8640.
- [11] W. S. Somers, J. Tang, G. D. Shaw, R. T. Camphausen, *Cell* **2000**, *103*, 467-479.
- [12] E. L. Berg, M. K. Robinson, O. Mansson, E. C. Butcher, J. L. Magnani, *J. Biol. Chem.* **1991**, *266*, 14869-14872.
- [13] R. M. Nelson, S. Dolich, A. Aruffo, O. Cecconi, M. P. Bevilacqua, *J. Clin. Invest.* **1993**, *91*, 1157-1166.
- [14] L. Poppe, G. S. Brown, J. S. Philo, P. V. Nikrad, B. H. Shah, *J. Am. Chem. Soc.* **1997**, *119*, 1727-1736.
- [15] K. E. Norman, G. P. Anderson, H. C. Kolb, K. Ley, B. Ernst, *Br. J. Pharmacol.* **1997**, *120*, P1-P1.
- [16] S. H. Li, D. K. Burns, J. M. Rumberger, D. H. Presky, V. L. Wilkinson, M. Anostario, Jr., B. A. Wolitzky, C. R. Norton, P. C. Familletti, K. J. Kim, et al., *J. Biol. Chem.* **1994**, *269*, 4431-4437.
- [17] R. C. Preston, S. Rabbani, F. P. C. Binder, S. Moes, J. L. Magnani, B. Ernst, *Glycobiology* **2014**, *24*, 592-601.
- [18] H. C. Kolb, B. Ernst, *Chem. Eur. J.* **1997**, *3*, 1571-1578.

Chapter 3. – The Lectin FimH

3.1. Introduction to FimH

3.1.1. Urinary Tract Infections

A urinary tract infection (UTI) is a bacterial infection, with involvement of the bladder, the urethra, or even the kidneys. The major cause of UTIs are gram-negative uropathogenic *Escherichia coli* (UPEC) that are responsible in 80 % of the cases, thus being the most prominent species¹⁻³. *Staphylococcus saprophyticus* has been identified as the second most common species in UTIs with a proportion of 10–20 % of the cases⁴. UTIs are a common infection disease, since each year millions of people suffer from such an infection, causing major costs in diagnosis and treatment. Approximately 50 % of all women endure at least one symptomatic UTI during their lifetime. Further risk factors are diabetes⁵, spinal cord injuries⁶, age (menopause⁷), suppressed immune system⁸, sexual activity and use of contraceptives (diaphragm⁹).

A cystitis, commonly referred to as a bladder infection, is a UTI that only affects the lower parts of the urinary tract, *i.e.* the urethra and the bladder. The typical symptoms are pain during micturition (dysuria), frequent urination (polyuria), and blood traces in the urine (haematuria). When left untreated, a cystitis may lead to a pyelonephritis, an infection of the upper urinary tract involving the kidneys. A pyelonephritis poses a serious health risk, as it may lead to sepsis, kidney failure, and finally death³. A major problem of UTIs is its reoccurrence. One in four women who have experienced an initial UTI will suffer a second UTI regardless of antibiotic treatment within six months¹⁰.

3.1.2. The Infection Cycle and Virulence Factors

UPECs undergo a well defined infection cycle within the host (Figure 3-1). The adherence of the bacteria to epithelial cells in the lower urinary tract is the first and one of the most crucial steps. This is mediated by the interaction of bacterial type 1 pili, which contains the lectin FimH, and oligosaccharides presented on the glycocalyx on bladder epithelial cells. The cell-surface glycoprotein uroplakin-Ia (UPIa), which contains high-mannosylated glycan structures, has been identified as an important ligand for UPECs^{11,12}. Once adhered to the urothelium, UPECs are protected from clearance by the flow of urine and can invade the host cells within hours¹³, finally leading to a UTI^{14,15}. The host defends against UPECs by exfoliation of the outer cell lining and by activating the immune response accompanied by a massive influx of leukocytes¹⁶. By forming biofilm-like intracellular bacterial communities (IBC), in which bacteria are in a quiescent state, and by penetrating deeper into the host

tissue, UPECs can evade the host immune response. Especially the IBCs pose a major challenge for the host immune system and for the treatment of UTIs. The bacteria therein are embedded in an extracellular polysaccharide matrix, which binds and inactivates almost any external agents¹⁷⁻¹⁹. UPECs can remain in IBCs for long periods after which the bacteria can re-emerge and infect neighboring epithelial cells, a mechanism associated with the high reoccurrence of UTIs.

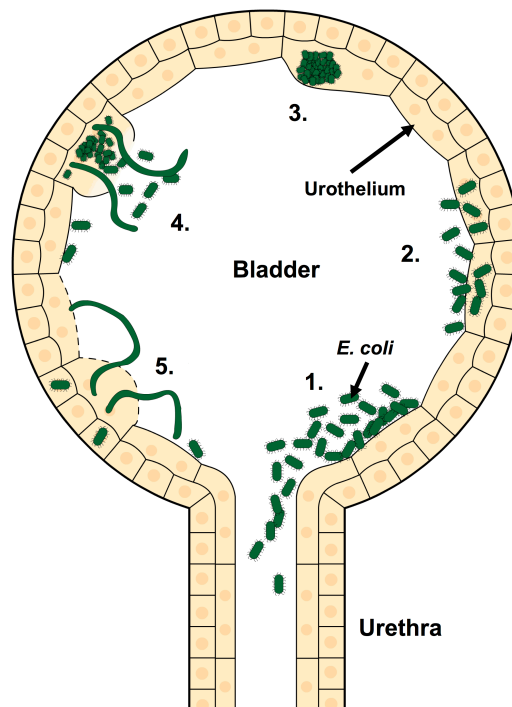


Figure 3-1. Infection cycle of uropathogenic *E. coli* (UPECs) in the lower urinary tract (adapted from²⁰). The urethra is the entry point of UPECs from which they colonize within the bladder. Initial adhesion (1) is mediated by type 1 pili binding to mannosylated glycoproteins (e.g. uroplakin Ia) presented on urothelial cells. UPECs consequently invade into the urothelium (2), in which they are able to form intracellular bacterial communities (3). Within these communities, the UPECs are protected from the host immune response and from antibiotics. UPECs are released as filamentous structures upon exfoliation of bladder epithelial cells (4) and are able either to penetrate deeper into the tissue or to infect surrounding urothelial cells (5).

UPECs have acquired a specific set of virulence factors to adhere to and invade host cells, and to form IBCs. These virulence factors are activated at the various stages of infection:

- **Adhesins**, which are required for attachment to epithelial cells¹⁴. The mannose specific type I fimbriae are involved in initial attachment; galabiose specific P pili are required for ascending infections.
- **Toxins** such as the lipoprotein α -haemolysin (HylA), which facilitate bacterial entry over mucosal barriers and damage immune cells²¹.
- **Siderophores** in the iron acquisition systems, which allows survival in nutrient poor environment²².
- Further virulence factors such as **proteases**²³, **lipopolysaccharides**, and the polysaccharide **capsule**, which are involved in IBC formation²².

3.1.3. Type I pili – Structure, Assembly and Expression

Type I pili are long, multimeric filamentous rods that protrude from the bacterial cell wall^{24,25}. Each pili is approximately 7 nm wide, 1-3 μ m in length and is composed of ~300-5000 of FimA subunits, followed by one of each FimF, FimG, and FimH that form the tip of the pili (see Figure 3-2). The subunits are connected by high-affinity, non-covalent interactions providing stability and rigidity^{26,27}. Each subunit has an incomplete immunoglobulin-like fold, with a missing C-terminal beta strand. During assembly in the periplasmic space, the subunits are bound to FimC, a chaperone protein which donates the missing beta strand and stabilizes the subunits^{28,29}. Upon delivery to FimD, an usher protein in the outer cell membrane, the subunit is released from FimC and is linked to the next subunit, which itself donates the missing beta strand. At the same time, FimD translocates the growing pilus rod through the outer cell membrane. A similar chaperone/usher pathway for pilus assembly is also used for the biosynthesis of other pili^{30,31}.

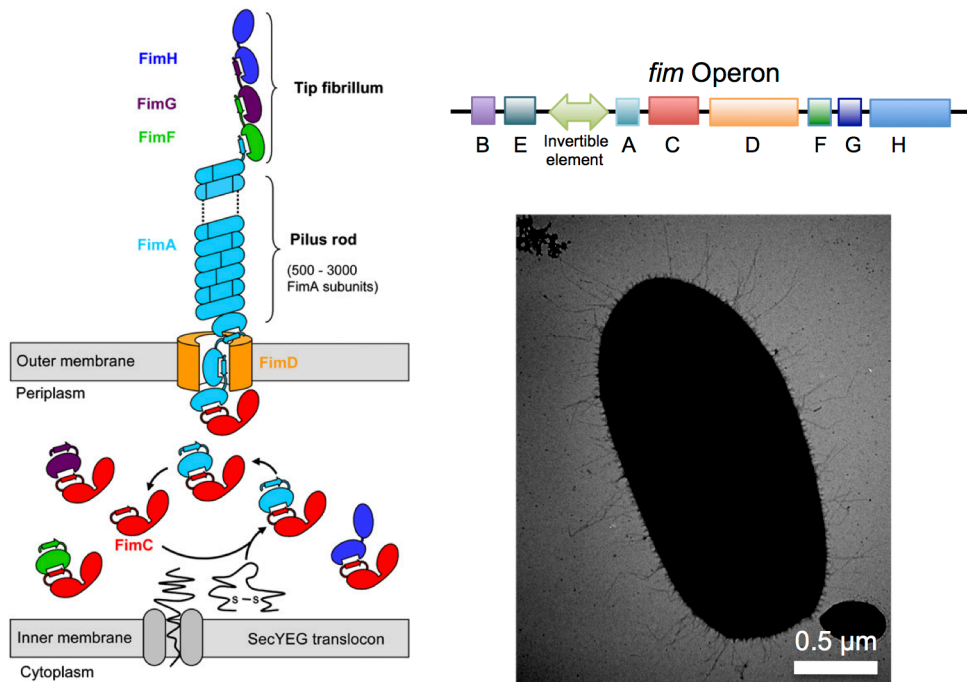


Figure 3-2. Left: Schematic representation of type 1 pili and their assembly²⁷. Pilus subunits are assembled via “donor strand complementation”, in which the immunoglobulin-like fold is completed by the insertion of an anti-parallel beta-sheet of the following subunit. The bacterial lectin FimH is located on the tip of each pilus. Right top: The *fim* operon encodes all type 1 pilus subunits as well as the chaperone FimC and the regulators FimB and FimE. The *fim* operon is regulated by an invertible element that regulates pilus expression. Right bottom: Electron micrograph of the UTI89 strain expressing type 1 pili (kindly provided by Dr. Meike Scharenberg and Deniz Eris, Institute of Molecular Pharmacy, University of Basel).

The genes for type I pili are located on the *fim* gene operon. Besides the genes encoding for the pili subunits, two regulatory genes are encoded (*fimB* and *fimE*). FimB and FimE are recombinases and are able to influence the direction of the invertible element, a 314 bp long region in the *fim* operon. In its off-state, the invertible element prevents mRNA transcription of the downstream *fimA*, *fimC*, *fimD*, *fimF*, *fimG*, and *fimH* genes. When switched to the on-state, transcription takes place and type I pili are expressed^{32,33}. The regulation of type I pili expression is also referred to as phase-variation. The two recombinases are influenced by external factors, *e.g.* pH, osmolarity, and motion. However, the exact mechanism is still elusive^{33,34}. While FimB is able to switch the invertible element in both directions (on-to-off and off-to-on), FimE can only promote expression silencing (on-to-off)^{35,36}. To prevent simultaneous expression of multiple pilus types, cross-talk between the various systems occur. For example, when the *fim* operon is in the on-state, the *pap* operon (encoding P pili) is suppressed³⁷. This strategy is important for the pathogenicity, as it allows specific pili

expression, depending on the surroundings of the bacteria. In the lower urinary tract, type I pili expression is preferred, while in the upper urinary tract, during the course of an ascending infection, P pili expression is beneficial.

3.1.4. The Bacterial Lectin FimH

The terminal FimH subunit on type I pili is responsible for the bacterial adherence, as it contains the mannose binding site^{24,38}. The 29 kDa protein with 279 amino acid residues consists, unlike the other Fim subunits, of two immunoglobulin-like domains: the lectin domain (FimH_{LD}, residues 1–156) at the N-terminus, which contains the CRD, and the C-terminal pilin domain (FimH_{PD}, residues 160–279), which connects FimH to the pilus rod (see Figure 3-3). The two domains are linked via three amino acid residues (amino acids 157–159). While the lectin domain alone is stable, the entire FimH protein consisting of both domains is not due to the missing beta-strand of the following subunit in the pilin domain. Therefore, full length FimH in solution is only stable in presence of FimC. The FimC-FimH complex was the first structure to be solved in 1999³⁸. Three years later Hung *et al.* reported the first FimC-FimH structure in presence of a mannoside ligand, which gave important insight into the binding site³⁹. Later, the structures of the FimH lectin domain alone⁴⁰ and in complex with oligomannose-3⁴¹ were published. An elegant crystal structure of FimH_{LD} and FimH_{PD} together with FimG, two FimF, and FimC subunits was reported by LeTrong *et al.* in 2010⁴². This structure revealed significant differences of the overall fold of FimH as discussed in section 3.1.7. Finally, several structures of the lectin domain with various mannoside ligands were recently published⁴³.

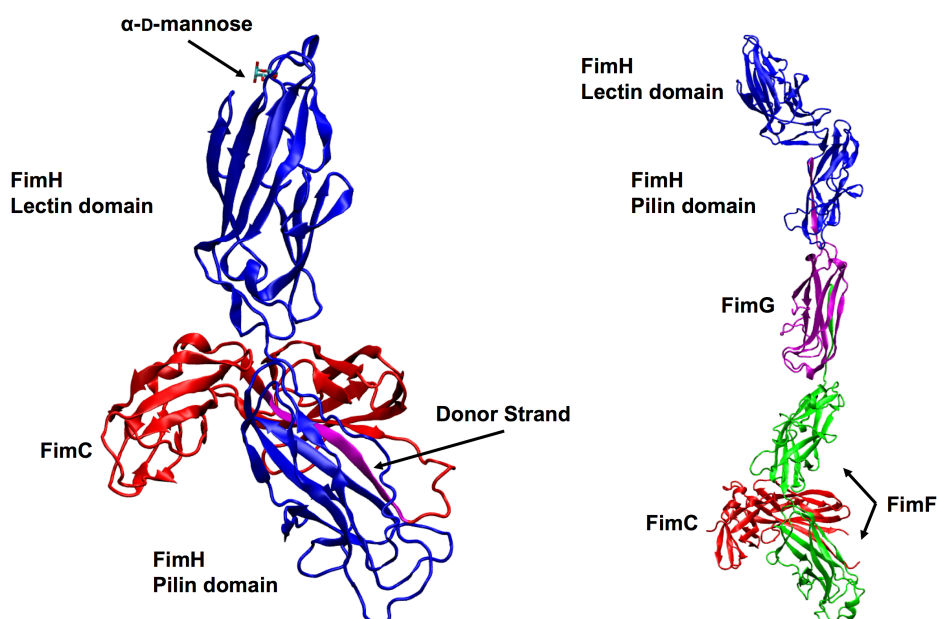


Figure 3-3. Left: Structure of FimH_{LD}-FimH_{PD}-FimC in complex with D-mannose (PDB code 1KLF)³⁹. FimH (blue) is wedged apart by FimC (red), which provides the donor strand (magenta). The carbohydrate binding site is located at the very tip of FimH. Right: Structure of the fimbrial tip (PDB code 3JWN)⁴² with both domains of FimH (blue), FimG (magenta), and FimC (red) wedged between two FimF domains (green).

3.1.5. Natural Ligands to FimH

As stated in section 3.1.1., type I pili dependent bacterial adhesion to the urothelium is mediated by the interaction of FimH to oligomannosides presented on cell surfaces. Therefore, FimH binds to any cell type presenting mannose, *e.g.* yeast cells (yeast mannan) and erythrocytes⁴⁴. On mammalian cells, mannose is presented on high-mannose *N*-linked glycans, consisting of five (Man₅-GlcNAc₂) to nine mannose (Man₉-GlcNAc₂) residues attached to two GlcNAc moieties (see Figure 3-4). FimH affinity strongly depends on the connectivity of the terminal mannose residues. For Man₉-GlcNAc₂, all terminal mannose moieties are α 1-2 linked, while for Man₅-GlcNAc₂, they are α 1-3 or α 1-6 linked. FimH shows strongest affinity towards α 1-3 linked mannose, followed by the α 1-6 and weakest binding to α 1-2 configuration⁴⁵.

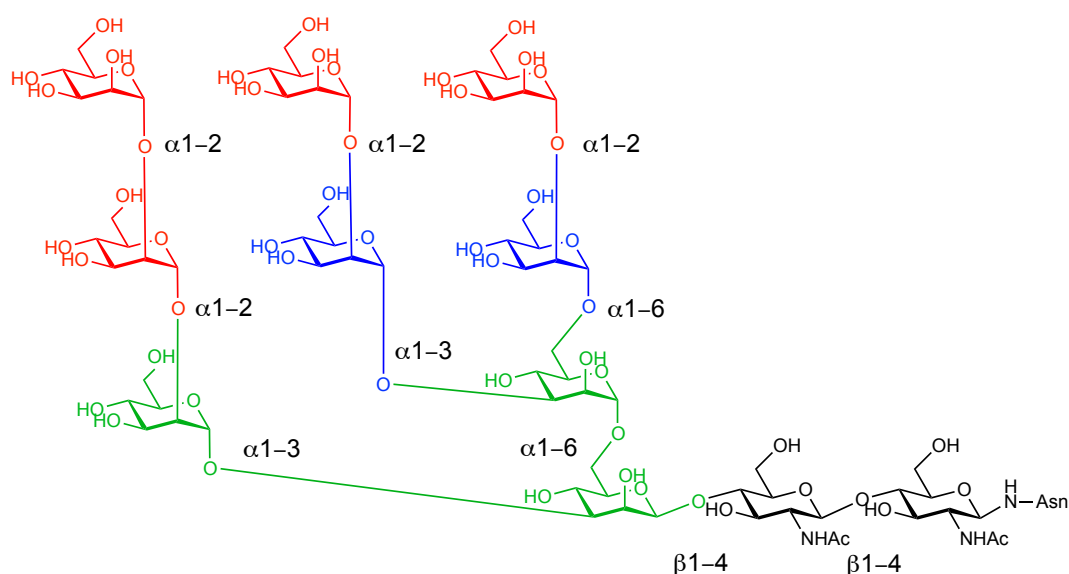


Figure 3-4. The natural ligand of FimH is presented on high-mannose *N*-linked glycans. These are presented on urothelial cells as part of the glycoprotein uroplakin Ia. While FimH binds weakly to full high-mannose *N*-linked glycans (Man₉-GlcNAc₂, red/blue/green), affinity towards truncated glycans (Man₅-GlcNAc₂, blue/green) is augmented.

Apart from mannose and mannosylated glycoproteins, FimH has been shown to bind non-carbohydrate structures as well. For example, the naturally occurring S62A mutation in the FimH_{LD} allows it to bind to type I and type IV collagen⁴⁶. Furthermore, FimH has been shown to bind fibronectin and peptide fragments thereof^{47,48}. The binding mechanism of these non-carbohydrate ligands has not been determined to date.

3.1.6. The Mannose Binding Site

From the crystal structure of FimH_{LD} with α -D-mannose (PDB code 1KLF)³⁹, the binding mode and the involved residues were identified. The mannose binding site consists of a deep, negatively charged pocket with a hydrophobic entrance. The residues Phe1, Asn46, Asp47, Asp54, Gln133, Asn135, Asp140, and Phe142 provide an extensive hydrogen bond network to the hydroxyl groups of mannose. All oxygen atoms of D-mannose are involved, except for the one at the anomeric center. The hydrophobic entrance, termed “tyrosine-gate”, is formed by Ile52, and the two aromatic residues Tyr48 and Tyr137. FimH has also been crystallized with *n*-butyl α -D-mannopyranoside (1UWF⁴⁰, see Figure 3-5), which sparked the development of α -D-mannopyranosides with hydrophobic aglycons as FimH antagonists (see Section 3.1.9.)⁴⁰.

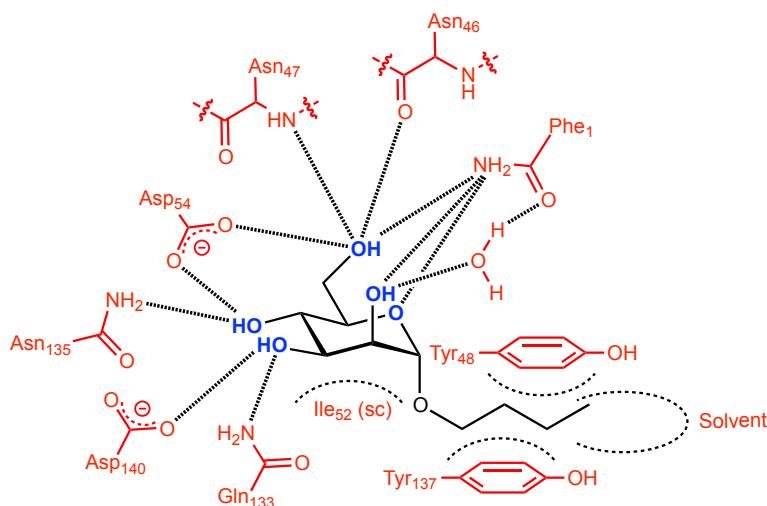


Figure 3-5. Schematic representation of the binding site of *n*-butyl α -D-mannopyranoside according to X-ray crystallography (PDB code 1UWF)⁴⁰. Relevant pharmacophores are marked in blue. Amino acid residues involved in binding are marked in red.

3.1.7. Catch-Bonds in FimH

Similar to the selectins, FimH also exhibits catch-bond binding behavior, *i.e.* an increased bond lifetime under tensile force⁴⁹. This has independently been observed under hydrodynamic force in flow chamber assays⁵⁰⁻⁵³ and in AFM studies⁵⁴. The structural mechanism by which the catch-bonding is regulated in FimH has been structurally elucidated⁴². FimH is able to adopt two distinct conformations, a high- and a low-affinity conformation. The low affinity conformation, which is the preferred conformation in absence of force or ligand, is stabilized by interdomain interactions between FimH_{LD} and FimH_{PD} with various loops (insertion-, swing-, and linker loop). These interdomain interactions twist the beta-sandwich fold of the FimH_{LD}, thus allosterically auto-inhibiting FimH by widening the binding pocket and not providing a suitable mannose binding site (see Figure 3-6). Upon transition into the high-affinity state by force along the FimH_{LD}-FimH_{PD} axis, interactions between the domains are disrupted, consequently relaxing and untwisting the FimH_{LD}. The interaction between the two domains for the low-affinity conformations are absolutely necessary to stabilize this conformation. Therefore, the FimH_{LD} expressed alone is always in the high-affinity state. This is also true for FimH-FimC co-expression, because FimC is wedged at the interdomain region, consequently forcing the two domains apart³⁸. Furthermore, mannose-based ligands can induce the switch from the low- to the high-affinity state, possibly by contracting the ligand binding site upon interaction⁴⁹.

It is assumed that the UPECs exploit the catch-bond function to have mobility in the bladder while avoiding clearance from the urinary tract under hydrodynamic force, as it occurs during micturition. Furthermore, lower affinity under static conditions prevents UPECs from clearance by natural soluble antagonists, such as the Tamm-Horsfall protein⁵⁵.

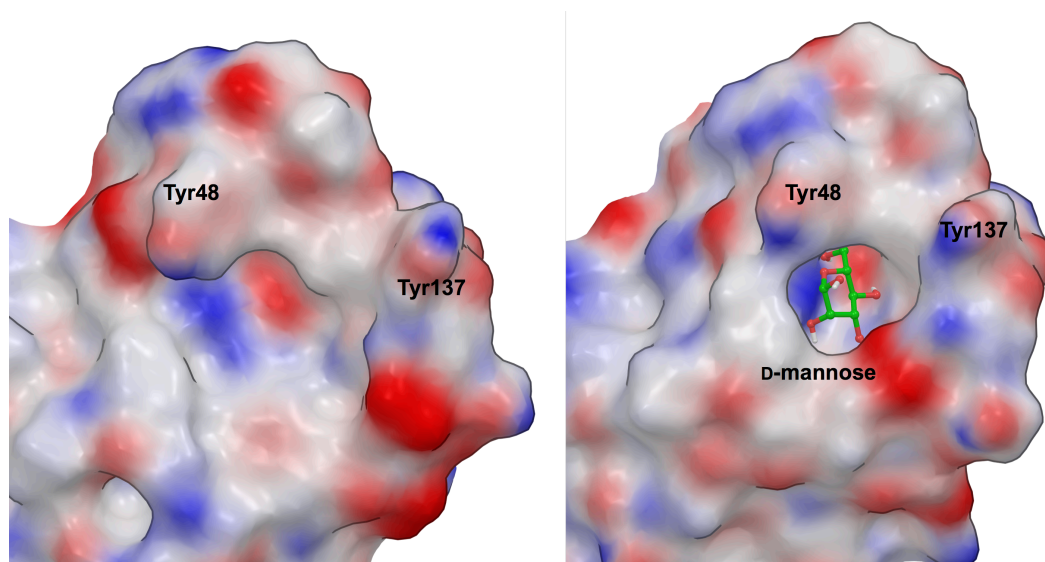


Figure 3-6. Comparison of the binding site in both states of FimH. Left: The low-affinity state, which is adopted without ligand and in absence of FimC (PDB code 3JWN)⁴². Right: D-mannose bound in the high-affinity state binding site (PDB code 1KLF)³⁹. In the high-affinity state, a deep pocket surrounding the ligand is formed, while being a shallow groove in the low-affinity state.

3.1.8. Variation of FimH

E. coli have a relatively high mutation rate of 1 in every 10^8 base pair during replication⁵⁶. As a consequence, UPECs genetically adapt to support optimal pathogenicity, a process, which also involves the *fimH* gene. The primary sequence of FimH was determined in several UPEC isolates from the intestine and from patients suffering from UTI^{57,58}. A number of single nucleotide point mutations were found, some of which alter the binding phenotype of FimH. Compared to non-pathogenic *E. coli* from the intestine, the mutations determined in UPECs usually lead to an enhanced affinity of FimH to mannose. Some of these mutations are found with a high frequency in UPECs, e.g. the A27V or the S62A mutation. Judging from the crystal structures of FimH, the functional implications of the mutations were rationalized: No mutation is located within the mannose binding pocket, which remains highly conserved throughout all investigated *E. coli* strains. All mutations are located in or near to the interdomain region of FimH_{LD} and FimH_{PD}. Therefore, it was assumed that these mutations influence the stability of the high- and low-affinity conformation, by disrupting the stabilizing interdomain bonds⁵³. Indeed, this was confirmed by Tchesnokova *et al.*, who

identified an antibody specific for a ligand induced binding site (LIBS)⁴⁹. The LIBS epitope is exposed only in the high-affinity state of FimH, while it is obstructed in the low-affinity conformation. By thorough analysis of many FimH mutants, the LIBS antibody binding was correlated with the mannose binding affinity. The higher the affinity towards mannose, the greater the binding of the LIBS antibody, which for the first time corroborated the allosteric link between ligand affinity and the conformational state.

The mutations leading to the destabilization of the allosteric autoinhibition by interfering with interdomain interactions, even under low shear or static conditions as found in the upper urinary tract, have evolved under positive selection due to their enhanced virulence as shown in *in vivo* studies⁵⁹.

3.1.9. FimH Antagonists

The first-line treatment for uncomplicated UTIs are a three-day antibiotic treatment with a combination of trimethoprim and sulfamethoxazole or with a fluoroquinolone (*e.g.* ciprofloxacin)^{60,61}. However, in clinical practice the emergence of bacterial resistance due to overuse of antibiotics poses a serious problem⁶². In a recently published study in which UPEC resistance in association with antibiotic treatment was monitored between 2000–2010, a dramatic increase in bacterial resistance was observed⁶³. To tackle the resistance problem, new strategies for UTI treatment are required. The inhibition of initial bacterial adherence by FimH antagonists is a promising approach. These anti-adhesive compounds must however fulfill a set of requirements besides being efficacious: they must be orally available, inexpensive, and safe. Furthermore, bacterial resistance mechanism should not be induced. Since FimH antagonists are neither bacteriolytic nor bacteriostatic, a reduced potential of bacterial resistance is expected.

Over recent years, numerous assays for the affinity determination of FimH antagonists have been reported. Initially, cell-based assays were employed in which the inhibition of aggregation of *E. coli* with guinea pig erythrocytes^{64,65} or yeast cells⁶⁶⁻⁶⁸ was determined. Furthermore, sophisticated cell-based assays using an immortalized urothelial cell line were developed⁶⁹. After successful cloning, expression and purification of FimH, target-based assays were developed. With competitive ELISA-based assays, high-throughput methods were established^{70,71}. Surface plasmon resonance studies yielded first equilibrium dissociation constants (K_D)⁴⁰, however, due to the assay setup, no association or dissociation constant could be determined. Isothermal titration calorimetry (ITC) provided further insights

into the thermodynamic aspects of FimH antagonist binding⁷²⁻⁷⁴. When comparing affinity data from various assays, variation in the assay setup must be considered. In target-based assays, usually the FimH_{LD} is employed, which is locked in the high-affinity state. In cell-based assays, however, FimH is can switch between the low- and high-affinity conformation. Furthermore, *E. coli* expressing different variants of FimH are employed, leading to a further diversification on the degree of high- and low-affinity FimH investigated in assays.

Since methyl α -D-mannopyranoside (**1**) has been shown to inhibit bacterial adhesion in the 1970s⁶⁶ with micro-^{40,71,75} to millimolar^{64,67} affinity, most drug discovery programs have focused on improving affinity by modifications on the aglycone portion at the anomeric center. Bouckaert and co-workers identified alkyl α -D-mannosides as potent FimH antagonists. A length of seven carbon atoms, as in *n*-heptyl α -D-mannopyranoside (**2**), proved to be the optimal length of the alkyl chain⁴⁰. Further hydrophobic aglycones were designed in order to reach the hydrophobic rim at the binding site entrance. Aromatic aglycons were introduced, *e.g.* *p*-nitrophenyl α -D-mannopyranoside⁶⁷ (**3**, *p*-NPMAN), which has a 30-fold increase in the relative inhibitory potential over methyl α -D-mannopyranoside⁷⁰. Extension of aromatic moieties lead to biphenyl aglycons modified with carboxylic acids (**4-9**)^{76,77} or other bulky substituents (**6-8**)^{73,77}. Indolinyphenyl derivatives (**10**) have proven to be beneficial for in oral availability, while maintaining affinity⁷⁸. Squaric acid derivatives (**11**) represent a class of FimH antagonists with affinities in the low nanomolar range⁷⁰. This was initially attributed to covalent binding of protein and ligand, a notion that was not confirmed⁷⁹. Further development was focused on the pharmacokinetic aspects of FimH antagonists to address oral availability. The efficacy of monovalent mannosides was determined in several *in vivo* mouse treatment studies^{41,78,80,81}. To reduce polarity and thus increase oral uptake by passive intestinal absorption, an ester prodrug approach was successfully employed⁷⁷. Indolinyphenyl proved to be successful candidates for oral application, since they were as effective as antibiotic treatment with ciprofloxacin⁷⁸. A representative set of the most important FimH antagonists are summarized in Figure 3-7.

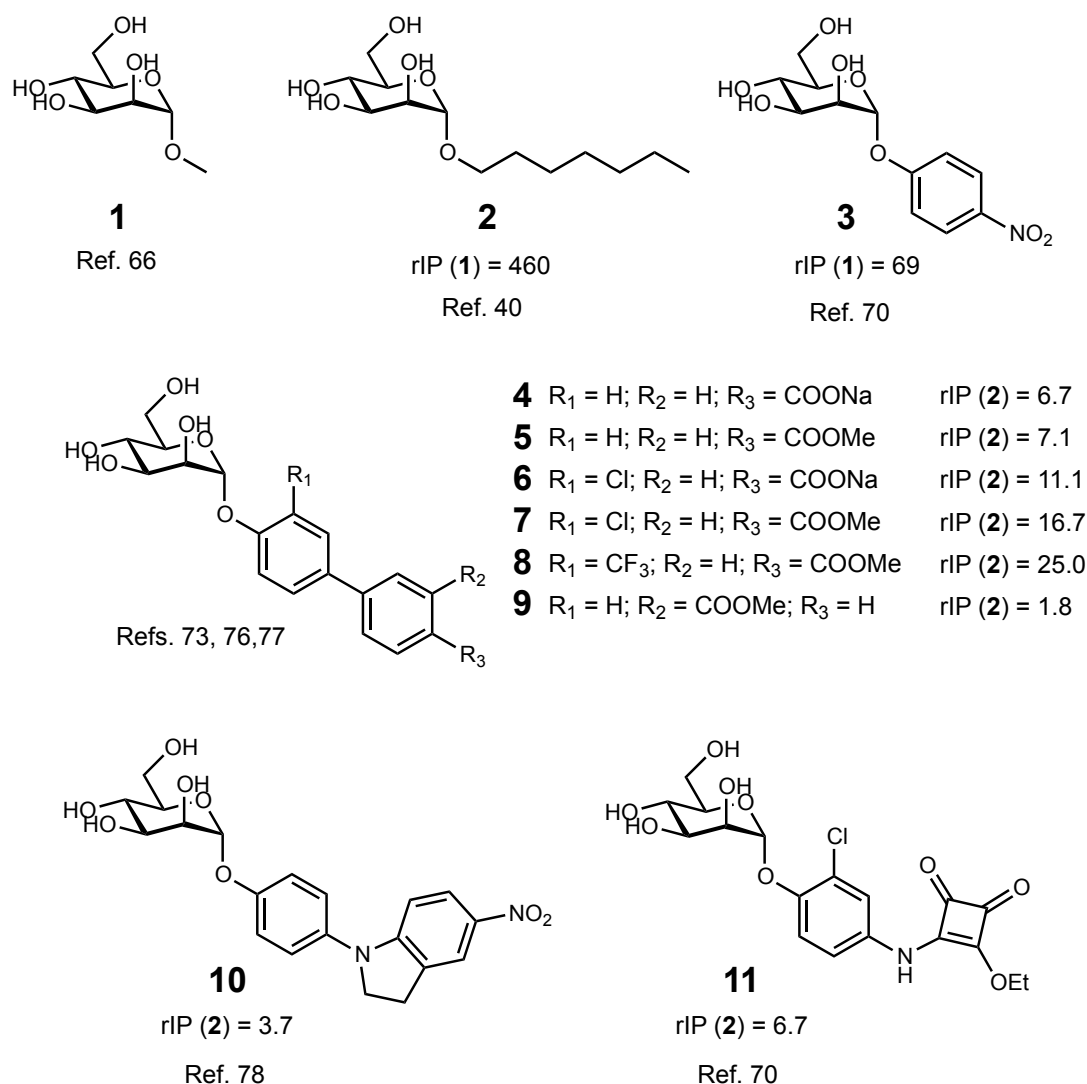


Figure 3-7. List of the most important FimH antagonists with their relative inhibitory potential (rIP). The rIPs were calculated either to compound **1** or **2** as noted in the brackets.

All mannosides with extended hydrophobic aglycons have improved affinity over α -D-mannose or methyl α -D-mannopyranoside due to hydrophobic interactions of the aglycone with the tyrosine gate. However, their binding mode can differ. For example, the aglycones of alkyl mannosides reside between Tyr48 and Tyr137⁴⁰. This binding mode is termed the “in-binding mode” (see Figure 3-8)⁸². The more bulky aromatic aglycons on the other hand, adopt a different conformation, in which π - π stacking only occurs exclusively with Tyr48 as shown in X-ray crystal structures⁷⁶. This binding mode is termed “out-binding mode”. However, recently published structures show aromatic aglycons, *e.g.* propenyl biphenyl α -D-mannopyranoside also binding in the in-docking mode⁴³. When discussing ligand poses in the crystal structures, one needs to take crystal packing effects into account. Binding site residues

form numerous crystal contacts in most published FimH structures, which might influence the location of certain residues and the ligand itself.

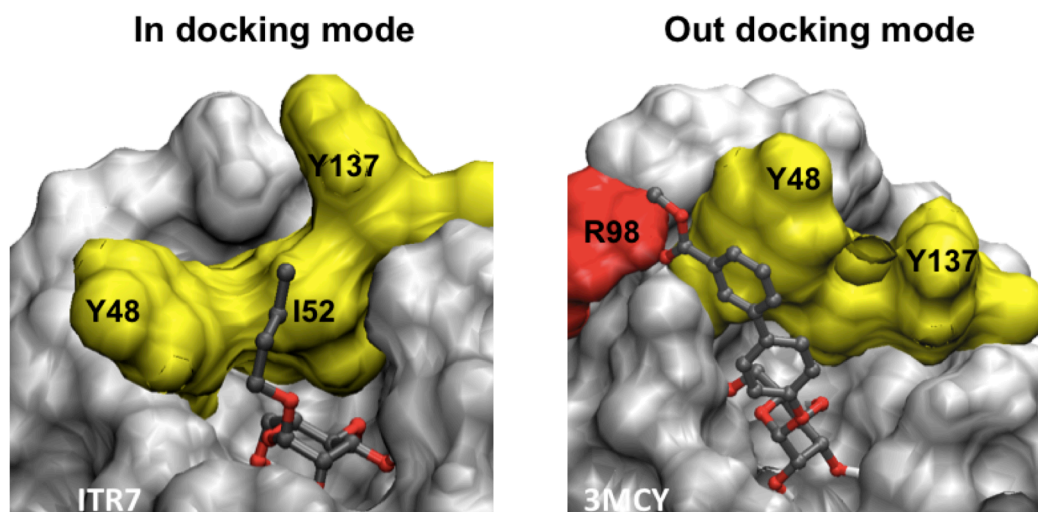


Figure 3-8. Two distinct binding modes of FimH antagonists. Left: The “in-binding mode” as observed in the FimH structure in complex with *n*-butyl α-D-mannopyranoside (PDB code 1TR7)⁴⁰. Right: The “out-binding mode” as observed with the biphenyl substituted α-D-mannopyranoside (**9**) (PDB code 3MCY)⁸³. By courtesy of Adam Zalewski, University of Basel.

Finally, multimeric FimH antagonists were developed in which mannose moieties are attached to a scaffold. Hartmann and Lindhorst have summarized the recent development of multimeric FimH antagonists in their review⁸⁴. While the affinity of mannose multimers is high due to avidity and their ability to bind multiple FimH molecules simultaneously, such antagonists are most likely not orally available due to their size and polarity.

3.1.10. Literature for Chapter 3.1.

1. Foxman, B. Recurring Urinary-Tract Infection - Incidence and Risk-Factors. *Am. J. Public Health* **80**, 331-333 (1990).
2. Foxman, B., Barlow, R., D'Arcy, H., Gillespie, B. & Sobel, J.D. Urinary tract infection: Self reported incidence and associated costs. *Ann. Epidemiol.* **10**, 509-515 (2000).
3. Ronald, A. The etiology of urinary tract infection: Traditional and emerging pathogens. *Am. J. Med.* **113**, 14S-19S (2002).
4. Kuroda, M., Yamashita, A., Hirakawa, H., Kumano, M., Morikawa, K., Higashide, M., Maruyama, A., Inose, Y., Matoba, K., Toh, H. et al. Whole genome sequence of *Staphylococcus saprophyticus* reveals the pathogenesis of uncomplicated urinary tract infection. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 13272-13277 (2005).
5. Patterson, J.E. & Andriole, V.T. Bacterial urinary tract infections in diabetes. *Infect. Dis. Clin. North Am.* **11**, 735-750 (1997).
6. Cardenas, D.D. & Hooton, T.M. Urinary-Tract Infection in Persons with Spinal-Cord Injury. *Arch. Phys. Med. Rehabil.* **76**, 272-280 (1995).
7. Foxman, B., Somsel, P., Tallman, P., Gillespie, B., Raz, R., Colodner, R., Kandula, D. & Sobel, J.D. Urinary tract infection among women aged 40 to 65: Behavioral and sexual risk factors. *J. Clin. Epidemiol.* **54**, 710-718 (2001).
8. Tolkoff-Rubin, N.E. & Rubin, R.H. Urinary tract infection in the immunocompromised host - Lessons from kidney transplantation and the AIDS epidemic. *Infect. Dis. Clin. North Am.* **11**, 707-717 (1997).
9. Hooton, T.M., Scholes, D., Hughes, J.P., Winter, C., Roberts, P.L., Stapleton, A.E., Stergachis, A. & Stamm, W.E. A prospective study of risk factors for symptomatic urinary tract infection in young women. *N. Engl. J. Med.* **335**, 468-474 (1996).
10. Franco, A.V.M. Recurrent urinary tract infections. *Best Pract Res Cl Ob* **19**, 861-873 (2005).
11. Zhou, G., Mo, W.J., Sebbel, P., Min, G.W., Neubert, T.A., Glockshuber, R., Wu, X.R., Sun, T.T. & Kong, X.P. Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: evidence from in vitro FimH binding. *J. Cell Sci.* **114**, 4095-4103 (2001).
12. Xie, B., Zhou, G., Chan, S.Y., Shapiro, E., Kong, X.P., Wu, X.R., Sun, T.T. & Costello, C.E. Distinct glycan structures of uroplakins Ia and Ib - Structural basis for the selective binding of FimH adhesin to uroplakin Ia. *J. Biol. Chem.* **281**, 14644-14653 (2006).
13. Martinez, J.J., Mulvey, M.A., Schilling, J.D., Pinkner, J.S. & Hultgren, S.J. Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO J.* **19**, 2803-2812 (2000).
14. Mulvey, M.A. Adhesion and entry of uropathogenic *Escherichia coli*. *Cell. Microbiol.* **4**, 257-271 (2002).
15. Wiles, T.J., Kulesus, R.R. & Mulvey, M.A. Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Exp. Mol. Pathol.* **85**, 11-19 (2008).
16. Mulvey, M.A., Lopez-Boado, Y.S., Wilson, C.L., Roth, R., Parks, W.C., Heuser, J. & Hultgren, S.J. Induction and evasion of host defenses by type 1-piliated uropathogenic *Escherichia coli*. *Science* **282**, 1494-1497 (1998).
17. Anderson, G.G., Dodson, K.W., Hooton, T.M. & Hultgren, S.J. Intracellular bacterial communities of uropathogenic *Escherichia coli* in urinary tract pathogenesis. *Trends Microbiol.* **12**, 424-430 (2004).

18. Anderson, G.G., Goller, C.C., Justice, S., Hultgren, S.J. & Seed, P.C. Polysaccharide Capsule and Sialic Acid-Mediated Regulation Promote Biofilm-Like Intracellular Bacterial Communities during Cystitis. *Infect. Immun.* **78**, 963-975 (2010).
19. Anderson, B.N., Ding, A.M., Nilsson, L.M., Kusuma, K., Tchesnokova, V., Vogel, V., Sokurenko, E.V. & Thomas, W.E. Weak rolling adhesion enhances bacterial surface colonization. *J. Bacteriol.* **189**, 1794-802 (2007).
20. Abgottspon, D. & Ernst, B. In vivo Evaluation of FimH Antagonists - A Novel Class of Antimicrobials for the Treatment of Urinary Tract Infection. *Chimia* **66**, 166-169 (2012).
21. Stanley, P., Koronakis, V. & Hughes, C. Acylation of Escherichia coli hemolysin: A unique protein lipidation mechanism underlying toxin function. *Microbiol. Mol. Biol. Rev.* **62**, 309-333 (1998).
22. Johnson, J.R. Microbial virulence determinants and the pathogenesis of urinary tract infection. *Infect. Dis. Clin. North Am.* **17**, 261-278 (2003).
23. Heimer, S.R., Rasko, D.A., Lockatell, C.V., Johnson, D.E. & Mobley, H.L.T. Autotransporter genes pic and tsh are associated with Escherichia coli strains that cause acute pyelonephritis and are expressed during urinary tract infection. *Infect. Immun.* **72**, 593-597 (2004).
24. Hahn, E., Wild, P., Hermanns, U., Sebbel, P., Glockshuber, R., Haner, M., Taschner, N., Burkhard, P., Aebi, U. & Muller, S.A. Exploring the 3D molecular architecture of Escherichia coli type 1 pili. *J. Mol. Biol.* **323**, 845-857 (2002).
25. Russell, P.W. & Orndorff, P.E. Lesions in 2 Escherichia-Coli Type-1 Pilus Genes Alter Pilus Number and Length without Affecting Receptor-Binding. *J. Bacteriol.* **174**, 5923-5935 (1992).
26. Nishiyama, M. & Glockshuber, R. The Outer Membrane Usher Guarantees the Formation of Functional Pili by Selectively Catalyzing Donor-Strand Exchange between Subunits That Are Adjacent in the Mature Pilus. *J. Mol. Biol.* **396**, 1-8 (2010).
27. Puorger, C., Eidam, O., Capitani, G., Erilov, D., Grutter, M.G. & Glockshuber, R. Infinite kinetic stability against dissociation of supramolecular protein complexes through donor strand complementation. *Structure* **16**, 631-642 (2008).
28. Hung, D.L. & Hultgren, S.J. Pilus biogenesis via the chaperone/usher pathway: An integration of structure and function. *J. Struct. Biol.* **124**, 201-220 (1998).
29. Sauer, F.G., Barnhart, M., Choudhury, D., Knights, S.D., Waksman, G. & Hultgren, S.J. Chaperone-assisted pilus assembly and bacterial attachment. *Curr. Opin. Struct. Biol.* **10**, 548-556 (2000).
30. Dodson, K.W., Jacobdubuisson, F., Striker, R.T. & Hultgren, S.J. Outer-Membrane Papc Molecular Usher Discriminately Recognizes Periplasmic Chaperone Pilus Subunit Complexes. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3670-3674 (1993).
31. Knight, S.D., Berglund, J. & Choudhury, D. Bacterial adhesins: structural studies reveal chaperone function and pilus biogenesis. *Curr. Opin. Chem. Biol.* **4**, 653-660 (2000).
32. Sohanpal, B.K., Kulasekara, H.D., Bonnen, A. & Blomfield, I.C. Orientational control of fimE expression in Escherichia coli. *Mol. Microbiol.* **42**, 483-494 (2001).
33. Gally, D.L., Rucker, T.J. & Blomfield, I.C. The Leucine-Responsive Regulatory Protein Binds to the Fim Switch to Control Phase Variation of Type-1 Fimbrial Expression in Escherichia-Coli K-12. *J. Bacteriol.* **176**, 5665-5672 (1994).
34. Tsai, K.W., Lai, H.T., Tsai, T.C., Wu, Y.C., Yang, Y.T., Chen, K.Y., Chen, C.M., Li, Y.S.J. & Chen, C.N. Difference in the regulation of IL-8 expression induced by

- uropathogenic *E. coli* between two kinds of urinary tract epithelial cells. *J. Biomed. Sci.* **16**(2009).
35. Klemm, P. The Fim Genes from *Escherichia-Coli* - Regulation of Fase Variation and Modulation of Expression. *J. Cell. Biochem.*, 138-138 (1986).
 36. Klemm, P. 2 Regulatory Fim Genes, Fimb and Fime, Control the Phase Variation of Type-1 Fimbriae in *Escherichia-Coli*. *EMBO J.* **5**, 1389-1393 (1986).
 37. Xia, Y., Gally, D., Forsman-Semb, K. & Uhlin, B.E. Regulatory cross-talk between adhesin operons in *Escherichia coli*: inhibition of type 1 fimbriae expression by the PapB protein. *EMBO J.* **19**, 1450-1457 (2000).
 38. Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S.J. & Knight, S.D. X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* **285**, 1061-1066 (1999).
 39. Hung, C.S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., Auguste, C.G., Strouse, R., Langermann, S., Waksman, G. et al. Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol. Microbiol.* **44**, 903-915 (2002).
 40. Bouckaert, J., Berglund, J., Schembri, M., De Genst, E., Cools, L., Wuhler, M., Hung, C.S., Pinkner, J., Slattegard, R., Zavialov, A. et al. Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.* **55**, 441-455 (2005).
 41. Wellens, A., Garofalo, C., Nguyen, H., Van Gerven, N., Slattegard, R., Hernalsteens, J.P., Wyns, L., Oscarson, S., De Greve, H., Hultgren, S. et al. Intervening with Urinary Tract Infections Using Anti-Adhesives Based on the Crystal Structure of the FimH-Oligomannose-3 Complex. *PLoS ONE* **3**(2008).
 42. Le Trong, I., Aprikian, P., Kidd, B.A., Forero-Shelton, M., Tchesnokova, V., Rajagopal, P., Rodriguez, V., Interlandi, G., Klevit, R., Vogel, V. et al. Structural Basis for Mechanical Force Regulation of the Adhesin FimH via Finger Trap-like beta Sheet Twisting. *Cell* **141**, 645-655 (2010).
 43. Wellens, A., Lahmann, M., Touaibia, M., Vaucher, J., Oscarson, S., Roy, R., Remaut, H. & Bouckaert, J. The Tyrosine Gate as a Potential Entropic Lever in the Receptor-Binding Site of the Bacterial Adhesin FimH. *Biochemistry (Mosc.)* **51**, 4790-4799 (2012).
 44. Giampapa, C.S., Abraham, S.N., Chiang, T.M. & Beachey, E.H. Isolation and Characterization of a Receptor for Type-1 Fimbriae of *Escherichia-Coli* from Guinea-Pig Erythrocytes. *J. Biol. Chem.* **263**, 5362-5367 (1988).
 45. Bouckaert, J., Mackenzie, J., de Paz, J.L., Chipwaza, B., Choudhury, D., Zavialov, A., Mannerstedt, K., Anderson, J., Pierard, D., Wyns, L. et al. The affinity of the FimH fimbrial adhesin is receptor-driven and quasi-independent of *Escherichia coli* pathotypes. *Mol. Microbiol.* **61**, 1556-1568 (2006).
 46. Pouttu, R., Puustinen, T., Virkola, R., Hacker, J., Klemm, P. & Korhonen, T.K. Amino acid residue Ala-62 in the FimH fimbrial adhesin is critical for the adhesiveness of meningitis-associated *Escherichia coli* to collagens. *Mol. Microbiol.* **31**, 1747-57 (1999).
 47. Sokurenko, E.V., Courtney, H.S., Ohman, D.E., Klemm, P. & Hasty, D.L. Fimh Family of Type-1 Fimbrial Adhesins - Functional-Heterogeneity Due to Minor Sequence Variations among Fimh Genes. *J. Bacteriol.* **176**, 748-755 (1994).
 48. Sokurenko, E.V., Courtney, H.S., Maslow, J., Siitonen, A. & Hasty, D.L. Quantitative Differences in Adhesiveness of Type-1 Fimbriated *Escherichia-Coli* Due to Structural Differences in Fimh Genes. *J. Bacteriol.* **177**, 3680-3686 (1995).

49. Tchesnokova, V., Aprikian, P., Yakovenko, O., Larock, C., Kidd, B., Vogel, V., Thomas, W. & Sokurenko, E. Integrin-like allosteric properties of the catch bond-forming FimH adhesin of *Escherichia coli*. *J. Biol. Chem.* **283**, 7823-7833 (2008).
50. Nilsson, L.M., Thomas, W.E., Sokurenko, E.V. & Vogel, V. Elevated shear stress protects *Escherichia coli* cells adhering to surfaces via catch bonds from detachment by soluble inhibitors. *Appl. Environ. Microbiol.* **72**, 3005-3010 (2006).
51. Nilsson, L.M., Thomas, W.E., Trintchina, E., Vogel, V. & Sokurenko, E.V. Catch bond-mediated adhesion without a shear threshold: trimannose versus monomannose interactions with the FimH adhesin of *Escherichia coli*. *J. Biol. Chem.* **281**, 16656-16663 (2006).
52. Nilsson, L.M., Yakovenko, O., Tchesnokova, V., Thomas, W.E., Schembri, M.A., Vogel, V., Klemm, P. & Sokurenko, E.V. The cysteine bond in the *Escherichia coli* FimH adhesin is critical for adhesion under flow conditions. *Mol. Microbiol.* **65**, 1158-1169 (2007).
53. Aprikian, P., Tchesnokova, V., Kidd, B., Yakovenko, O., Yarov-Yarovoy, V., Trintchina, E., Vogel, V., Thomas, W. & Sokurenko, E. Interdomain interaction in the FimH adhesin of *Escherichia coli* regulates the affinity to mannose. *J. Biol. Chem.* **282**, 23437-46 (2007).
54. Yakovenko, O., Tchesnokova, V., Aprikian, P., Forero, M., Vogel, V., Sokurenko, E. & Thomas, W. Mechanical force activates an adhesion protein through allosteric regulation. *Biophys. J.*, 348A-349A (2007).
55. Saemann, M.D., Weichhart, T., Horl, W.H. & Zlabinger, G.J. Tamm-Horsfall protein: a multilayered defence molecule against urinary tract infection. *Eur. J. Clin. Invest.* **35**, 227-235 (2005).
56. Drake, J.W. The distribution of rates of spontaneous mutation over viruses, prokaryotes, and eukaryotes. *Ann Ny Acad Sci* **870**, 100-107 (1999).
57. Sokurenko, E.V., Chesnokova, V., Dykhuizen, D.E., Ofek, I., Wu, X.R., Krogfelt, K.A., Struve, C., Schembri, M.A. & Hasty, D.L. Pathogenic adaptation of *Escherichia coli* by natural variation of the FimH adhesin. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 8922-8926 (1998).
58. Sokurenko, E.V., Feldgarden, M., Trintchina, E., Weissman, S.J., Avagyan, S., Chattopadhyay, S., Johnson, J.R. & Dykhuizen, D.E. Selection footprint in the FimH adhesin shows pathoadaptive niche differentiation in *Escherichia coli*. *Mol. Biol. Evol.* **21**, 1373-1383 (2004).
59. Chen, S.L., Hung, C.S., Pinkner, J.S., Walker, J.N., Cusumano, C.K., Li, Z.L., Bouckaert, J., Gordon, J.I. & Hultgren, S.J. Positive selection identifies an in vivo role for FimH during urinary tract infection in addition to mannose binding. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 22439-22444 (2009).
60. Nicolle, L., Anderson, P.A.M., Conly, J., Mainprize, T.C., Meuser, J., Nickel, J.C., Senikas, V.M. & Zhanel, G.G. Uncomplicated urinary tract infection in women - Current practice and the effect of antibiotic resistance on empiric treatment. *Can. Fam. Physician* **52**, 612-618 (2006).
61. Naber, K.G. Which fluoroquinolones are suitable for the treatment of urinary tract infections? *Int. J. Antimicrob. Agents* **17**, 331-341 (2001).
62. Levy, S.B. Antibiotic resistance - the problem intensifies. *Adv Drug Deliver Rev* **57**, 1446-1450 (2005).
63. Sanchez, G.V., Master, R.N., Karlowisky, J.A. & Bordon, J.M. In Vitro Antimicrobial Resistance of Urinary *Escherichia coli* Isolates among US Outpatients from 2000 to 2010. *Antimicrob. Agents Chemother.* **56**, 2181-2183 (2012).

64. Lindhorst, T.K., Kieburg, C. & Krallmann-Wenzel, U. Inhibition of the type 1 fimbriae-mediated adhesion of *Escherichia coli* to erythrocytes by multiantennary alpha-mannosyl clusters: The effect of multivalency. *Glycoconj. J.* **15**, 605-613 (1998).
65. Abgottspon, D., Rolli, G., Hosch, L., Steinhuber, A., Jiang, X.H., Schwardt, O., Cutting, B., Smiesko, M., Jenal, U., Ernst, B. et al. Development of an aggregation assay to screen FimH antagonists. *J. Microbiol. Methods* **82**, 249-255 (2010).
66. Ofek, I., Mirelman, D. & Sharon, N. Adherence of *Escherichia-Coli* to Human Mucosal Cells Mediated by Mannose Receptors. *Nature* **265**, 623-625 (1977).
67. Firon, N., Ashkenazi, S., Mirelman, D., Ofek, I. & Sharon, N. Aromatic Alpha-Glycosides of Mannose Are Powerful Inhibitors of the Adherence of Type-1 Fimbriated *Escherichia-Coli* to Yeast and Intestinal Epithelial-Cells. *Infect. Immun.* **55**, 472-476 (1987).
68. Firon, N., Ofek, I. & Sharon, N. Carbohydrate Specificity of the Surface Lectins of *Escherichia-Coli*, *Klebsiella-Pneumoniae*, and *Salmonella-Typhimurium*. *Carbohydr. Res.* **120**, 235-249 (1983).
69. Scharenberg, M., Abgottspon, D., Cicek, E., Jiang, X.H., Schwardt, O., Rabbani, S. & Ernst, B. A Flow Cytometry-Based Assay for Screening FimH Antagonists. *Assay Drug Dev. Technol.* **9**, 455-464 (2011).
70. Sperling, O., Dubber, M. & Lindhorst, T.K. Functionalization of oligosaccharide mimetics and multimerization using squaric diester-mediated coupling. *Carbohydr. Res.* **342**, 696-703 (2007).
71. Rabbani, S., Jiang, X.H., Schwardt, O. & Ernst, B. Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.* **407**, 188-195 (2010).
72. Durka, M., Buffet, K., Iehl, J., Holler, M., Nierengarten, J.F., Taganna, J., Bouckaert, J. & Vincent, S.P. The functional valency of dodecamannosylated fullerenes with *Escherichia coli* FimH-towards novel bacterial antiadhesives. *Chem. Commun. (Camb.)* **47**, 1321-1323 (2011).
73. Pang, L., Kleeb, S., Lemme, K., Rabbani, S., Scharenberg, M., Zalewski, A., Schadler, F., Schwardt, O. & Ernst, B. FimH Antagonists: Structure-Activity and Structure-Property Relationships for Biphenyl alpha-D-Mannopyranosides. *Chemmedchem* **7**, 1404-1422 (2012).
74. Pang, L.J., Kleeb, S., Lemme, K., Rabbani, S., Scharenberg, M., Zalewski, A., Schadler, F., Schwardt, O. & Ernst, B. FimH Antagonists: Structure-Activity and Structure-Property Relationships for Biphenyl a-D-Mannopyranosides. *Chemmedchem* **7**, 1404-1422 (2012).
75. Horst, A.K., Kotter, S., Lindhorst, T.K., Ludwig, A., Brandt, E. & Wagener, C. Binding inhibition of type 1 fimbriae to human granulocytes: a flow cytometric inhibition assay using trivalent cluster mannosides. *Med. Microbiol. Immunol. (Berl.)* **190**, 145-149 (2001).
76. Han, Z.F., Pinkner, J.S., Ford, B., Obermann, R., Nolan, W., Wildman, S.A., Hobbs, D., Ellenberger, T., Cusumano, C.K., Hultgren, S.J. et al. Structure-Based Drug Design and Optimization of Mannoside Bacterial FimH Antagonists. *J. Med. Chem.* **53**, 4779-4792 (2010).
77. Klein, T., Abgottspon, D., Wittwer, M., Rabbani, S., Herold, J., Jiang, X.H., Kleeb, S., Luthi, C., Scharenberg, M., Bezencon, J. et al. FimH Antagonists for the Oral Treatment of Urinary Tract Infections: From Design and Synthesis to in Vitro and in Vivo Evaluation. *J. Med. Chem.* **53**, 8627-8641 (2010).

78. Jiang, X.H., Abgottspon, D., Kleeb, S., Rabbani, S., Scharenberg, M., Wittwer, M., Haug, M., Schwardt, O. & Ernst, B. Antiadhesion Therapy for Urinary Tract Infections-A Balanced PK/PD Profile Proved To Be Key for Success. *J. Med. Chem.* **55**, 4700-4713 (2012).
79. Grabosch, C., Hartmann, M., Schmidt-Lassen, J. & Lindhorst, T.K. Squaric Acid Monoamide Mannosides as Ligands for the Bacterial Lectin FimH: Covalent Inhibition or Not? *Chembiochem* **12**, 1066-1074 (2011).
80. Aronson, M., Medalia, O., Schori, L., Mirelman, D., Sharon, N. & Ofek, I. Prevention of Colonization of the Urinary-Tract of Mice with Escherichia-Coli by Blocking of Bacterial Adherence with Methyl Alpha-D-Mannopyranoside. *J. Infect. Dis.* **139**, 329-332 (1979).
81. Eden, C.S., Freter, R., Hagberg, L., Hull, R., Hull, S., Leffler, H. & Schoolnik, G. Inhibition of Experimental Ascending Urinary-Tract Infection by an Epithelial Cell-Surface Receptor Analog. *Nature* **298**, 560-562 (1982).
82. Schwardt, O., Rabbani, S., Hartmann, M., Abgottspon, D., Wittwer, M., Kleeb, S., Zalewski, A., Smiesko, M., Cutting, B. & Ernst, B. Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.* **19**, 6454-6473 (2011).
83. Han, Z.F., Pinkner, J.S., Ford, B., Chorell, E., Crowley, J.M., Cusumano, C.K., Campbell, S., Henderson, J.P., Hultgren, S.J. & Janetka, J.W. Lead Optimization Studies on FimH Antagonists: Discovery of Potent and Orally Bioavailable Ortho-Substituted Biphenyl Mannosides. *J. Med. Chem.* **55**, 3945-3959 (2012).
84. Hartmann, M. & Lindhorst, T.K. The Bacterial Lectin FimH, a Target for Drug Discovery - Carbohydrate Inhibitors of Type 1 Fimbriae-Mediated Bacterial Adhesion. *Eur J Org Chem*, 3583-3609 (2011).

Chapter 3.2. – Manuscript 4

FimH Antagonists - Bioisosteres to Improve the *in vitro* and *in vivo* PK/PD Profile.

Preliminary manuscript formatted according to the requirements of *ChemBioChem*.

Contributions

S. Kleeb	<ul style="list-style-type: none">• Evaluation of PK properties.
L. Pang	<ul style="list-style-type: none">• Compound synthesis.
K. Mayer	<ul style="list-style-type: none">• Compound synthesis.
D. Eriş	<ul style="list-style-type: none">• Affinity measurements.
A. Sigl	<ul style="list-style-type: none">• <i>In vivo</i> PK experiments.• <i>In vivo</i> infection experiments.
R.C. Preston	<ul style="list-style-type: none">• Thermodynamic profiling by ITC and data analysis.• Protein-ligand co-crystallography, data acquisition, and data analysis.
P. Zihlmann	<ul style="list-style-type: none">• Aided in thermodynamic profiling by ITC and data analysis.
T. Sharpe	<ul style="list-style-type: none">• Supervision of and provided expertise in affinity measurements.
R. P. Jakob	<ul style="list-style-type: none">• Aided in protein-ligand co-crystallography, data acquisition, and data analysis.
D. Abgottspon	<ul style="list-style-type: none">• <i>In vivo</i> PK experiments.• <i>In vivo</i> infection experiments.
A. S. Hutter	<ul style="list-style-type: none">• Aided in affinity measurements.

**Contributions
continued**

M. Scharenberg	<ul style="list-style-type: none">• Supervision of and provided expertise in affinity measurements.
X Jiang	<ul style="list-style-type: none">• Compound synthesis.
G. Navarra	<ul style="list-style-type: none">• Compound synthesis.
S. Rabbani	<ul style="list-style-type: none">• Protein expression and purification.
Martin Smiesko	<ul style="list-style-type: none">• Calculation of desolvation energies.
N. Lüdín	<ul style="list-style-type: none">• Aided in PK profiling.
J. Bezençon	<ul style="list-style-type: none">• Physicochemical properties and <i>in vitro</i> pharmacokinetics.
O. Schwardt	<ul style="list-style-type: none">• Compound synthesis.
T. Maier	<ul style="list-style-type: none">• Supervision of protein-ligand co-crystallography, data acquisition, and data analysis.

FimH Antagonists - Bioisosteres to Improve the *in vitro* and *in vivo* PK/PD

Profile

Simon Kleeb,^{a)}# Lijuan Pang,^{a)}# Katharina Mayer,^{a)}# Deniz Eris,^{a)}# Anja Sigl,^{a)}# Roland C. Preston,^{a)} Pascal Zihlmann,^{a)} Timothy Sharpe,^{c)} Roman P. Jakob,^{b)} Daniela Abgottspon,^{a)} Aline S. Hutter,^{a)} Meike Scharenberg,^{a)} Xiaohua Jiang,^{a)} Giulio Navarra,^{a)} Said Rabbani,^{a)} Martin Smiesko,^{a)} Nathalie Lüdin,^{a)} Jacqueline Bezençon,^{a)} Oliver Schwardt,^{a)} Timm Maier,^{b)} Beat Ernst^{a)*}

Equally contributed

^{a)} Institute of Molecular Pharmacy, Pharmacenter, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland

^{b)} Structural Biology, Biocenter, University of Basel, Klingelbergstrasse 70, CH-4056 Basel

^{c)} Biophysical Facility, Biocenter, University of Basel, Klingelbergstrasse 70, CH-4056 Basel

* To whom correspondence should be addressed: Prof. Dr. Beat Ernst, Institute of Molecular Pharmacy, Pharmacenter, University of Basel, Klingelbergstrasse 50, CH-4056 Basel, Switzerland; Tel: +41 61 267 15 51, Fax: +41 61 267 15 52; E-mail: beat.ernst@unibas.ch

ABBREVIATIONS:

ΔH , change in enthalpy; ΔS , change in entropy; AUC, Area under the curve; BSA, bovine serum albumin; C_{\max} , maximal concentration; Caco-2 cells, colorectal adenocarcinoma cells; CFU, colony forming units; CL_{tot} , total clearance; CRD, carbohydrate recognition domain; C_0 , initial concentration; FITC, fluorescein isothiocyanate; FP, fluorescence polarization; ITC, isothermal titration calorimetry; i.v., intravenous; K_D , dissociation constant; MAC_{90} , minimal anti-adhesion concentration to inhibit 90% adhesion; PAMPA, parallel artificial membrane permeation assay; P_{app} , apparent permeability; PD, pharmacodynamics; P_e , effective

permeability; PK, pharmacokinetics; po, per os; s.c. subcutaneous; UPEC, uropathogenic *Escherichia coli*; UTI, urinary tract infection; V_z , volume of distribution in terminal phase.

ABSTRACT

Urinary tract infections (UTIs), predominantly caused by uropathogenic *Escherichia coli* (UPEC) belong to the most prevalent infectious diseases worldwide. The attachment of UPEC to host cells is mediated by FimH, a mannose-binding adhesin at the tip of bacterial type 1 pili. To date, UTIs are mainly treated with antibiotics, leading to the ubiquitous problem of increasing resistance against most of the currently available antimicrobials. Therefore, new treatment strategies are urgently needed. Here, we describe the development of an orally available FimH antagonist. Starting from the carboxylate substituted biphenyl α -D-mannoside **9** affinity as well as the relevant pharmacokinetic parameters (solubility, permeability, renal excretion) were substantially improved by a bioisosteric approach. With 3'-chloro-4'-(α -D-mannopyranosyloxy)-biphenyl-4-carbonitrile (**10j**) a FimH antagonist with an optimal *in vitro* PK/PD profile was identified. Orally applied, **10j** was effective in a mouse model of UTI by reducing the bacterial load in the bladder by over 1000-fold.

INTRODUCTION

Urinary tract infection (UTI) is one of the most frequent infectious diseases worldwide and affects millions of people every year.¹ In more than 70% of the reported cases, uropathogenic *Escherichia coli* (UPEC) is the causal pathogen.² Acute, uncomplicated lower urinary tract infection, commonly referred to as cystitis, requires an antibiotic treatment for symptom relief (i.e. reduction of dysuria, frequent and urgent urination, bacteriuria, pyuria) and for prevention of more devastating or even life threatening complications like pyelonephritis and urosepsis.^{3,4} However, the repeated use of antibacterial chemotherapeutics provokes antimicrobial resistance leading to treatment failure.⁵ Hence, a new approach for the prevention and treatment of UTI with orally applicable therapeutics is urgently needed.⁶

UPEC undergo a well-defined infection cycle within the host.⁷ The key step in pathogenesis is bacterial adhesion to the epithelial cells in the lower urinary tract.⁸ This interaction prevents UPEC from clearance by the bulk flow of urine and enables the bacteria to colonize the epithelial cells. The adhesion is mediated by the virulence factor FimH located at the tip of bacterial type 1 pili.^{9,10} FimH consists of two immunoglobulin-like domains: the N-terminal lectin domain and – connected by a short linker – the C-terminal pilin domain.¹¹ The lectin domain encloses the carbohydrate recognition domain (CRD) that binds to the oligomannosides of the glycoprotein uroplakin Ia on the epithelial cell surface.¹² The pilin domain anchors the adhesin to the pilus and regulates the switch between two conformational states of the CRD with high and low affinity for mannosides, respectively.

More than three decades ago, Sharon and co-workers described various oligomannosides and aryl α -D-mannosides as potential antagonists of the FimH-mediated bacterial adhesion.^{13,14}

However, only weak interactions in the milli- to micromolar range were observed. In recent years, several high-affinity monovalent mannose-based FimH antagonists with various aglycones like *n*-alkyl,¹⁵ phenyl,¹⁶ dioxocyclobutenylaminophenyl,¹⁷ umbelliferyl,¹⁶ biphenyl,¹⁸⁻²² indol(in)ylphenyl,²³ triazolyl²⁴ and thiazolylamino²⁵ have been reported. In addition, different multivalent presentations of the mannose have been synthesized²⁶⁻³² and a heptavalent presentation of *n*-heptyl α -D-mannoside (**1**) tethered to β -cyclodextrin proved to be highly effective when applied together with the UTI89 bacterial strain through a catheter into the bladder of C3H/HeN mice.³² Importantly, adverse side effects resulting from non-selective binding of FimH antagonists - they are all α -D-mannopyranosides - to mannose receptors of the human host system have recently been ruled out.³³

The high affinities of the monovalent α -D-mannopyranosides are based on optimal interactions with the main structural features of the CRD.³⁴⁻³⁷ First, the mannose binding pocket accommodating the mannose moiety by means of an extended hydrogen bond network and, second, the entrance to the binding site composed of three hydrophobic amino acids (Tyr48, Tyr137, and Ile52) and therefore referred to as ‘tyrosine gate’ hosting aliphatic and aromatic aglycones. As an example, *n*-heptyl α -D-mannopyranoside (**1**) exhibits nanomolar affinity due to hydrophobic contacts of the alkyl aglycone with the hydrophobic residues of the tyrosine gate.¹⁵ Furthermore, aromatic aglycones, such as present in mannosides **2** & **3** (Figure 1), provide strong π - π stacking interactions with the tyrosine gate. This interaction is further favored by the addition of an electron withdrawing substituent on the terminal ring of the biaryl portion (\rightarrow **4**).^{18,19}

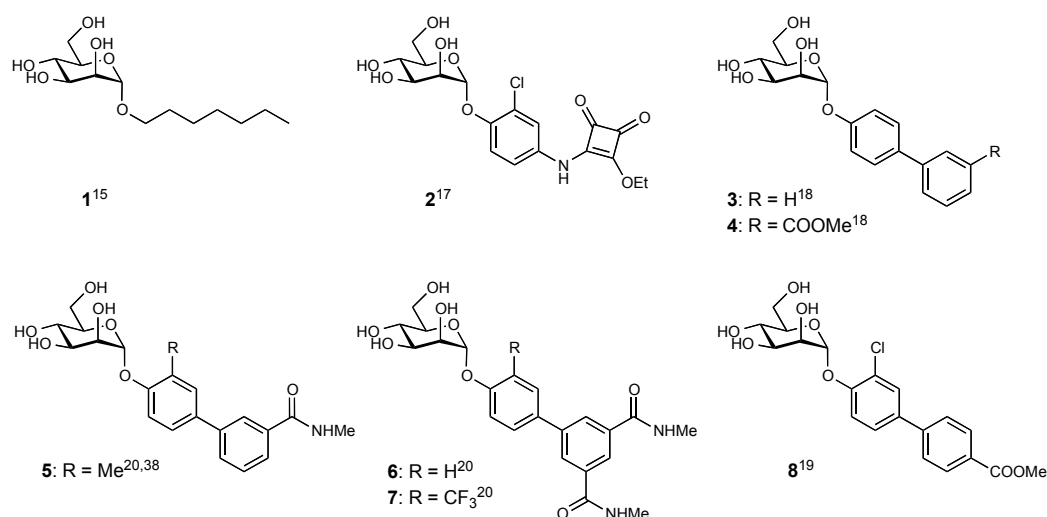


Figure 1. Monovalent FimH antagonists **1-4** acting as reference compounds and **5-8** which have been orally explored in *in vivo* disease models.

Recent *in vivo* PK studies in mice proved the high potential of the biphenyl α -D-mannosides **5-8** for an oral treatment, although high doses (≥ 50 mg/kg) were necessary to achieve the minimal concentrations required for the anti-adhesive effect in the urinary bladder.¹⁹⁻²¹ Moreover, the therapeutic effect could only be maintained for a few hours, i.e. four hours for a po (per os) single-dose application of **7** (50 mg/kg), because of rapid elimination by glomerular filtration and low reabsorption from the primary urine in the renal tubules.¹⁹

To date, the physicochemical properties affecting the rate of renal excretion, i.e. lipophilicity and plasma protein binding (PPB), or metabolic liabilities promoting non-renal elimination pathways have been barely investigated for FimH antagonists. The goal of the present study was to optimize the biphenyl α -D-mannoside with respect to oral bioavailability and renal excretion. Starting from antagonist **9**¹⁹ (Figure 2), we synthesized new biphenyl derivatives, characterized their affinity to the CRD, structurally investigated their binding mode, and determined physicochemical and pharmacokinetic parameters predictive for intestinal

absorption and renal elimination. Furthermore, we determined *in vivo* PK (pharmacokinetics) of the most promising new antagonists in a mouse model. After oral administration, the compound with the best PK profile proved effective in reducing the bacterial loads upon bladder infection in a mouse model of UTI.

RESULTS AND DISCUSSION

As previously reported, the carboxylate substituent present in the biphenyl mannoside **9** – its electron withdrawing potential being essential for an enhanced drug target interaction – strongly decreases the lipophilicity of the antagonist ($\log D_{7.4} < -1.5^{19}$) in comparison to the *n*-heptyl (\rightarrow **1**, $\log P = 1.7^{19}$) or the unsubstituted biphenyl aglycone (\rightarrow **3**, $\log P = 2.1^{22}$). Since low lipophilicity is a major reason for low intestinal absorption and rapid renal excretion of the systemically available antagonist,^{19,23} we aspired to improve oral bioavailability as well as renal excretion by replacing the carboxylate in **9** with various bioisosteric groups³⁹ (Figure 2).

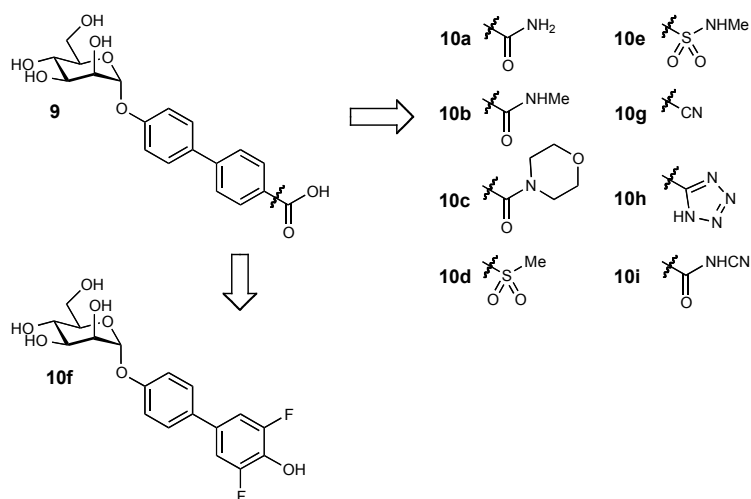
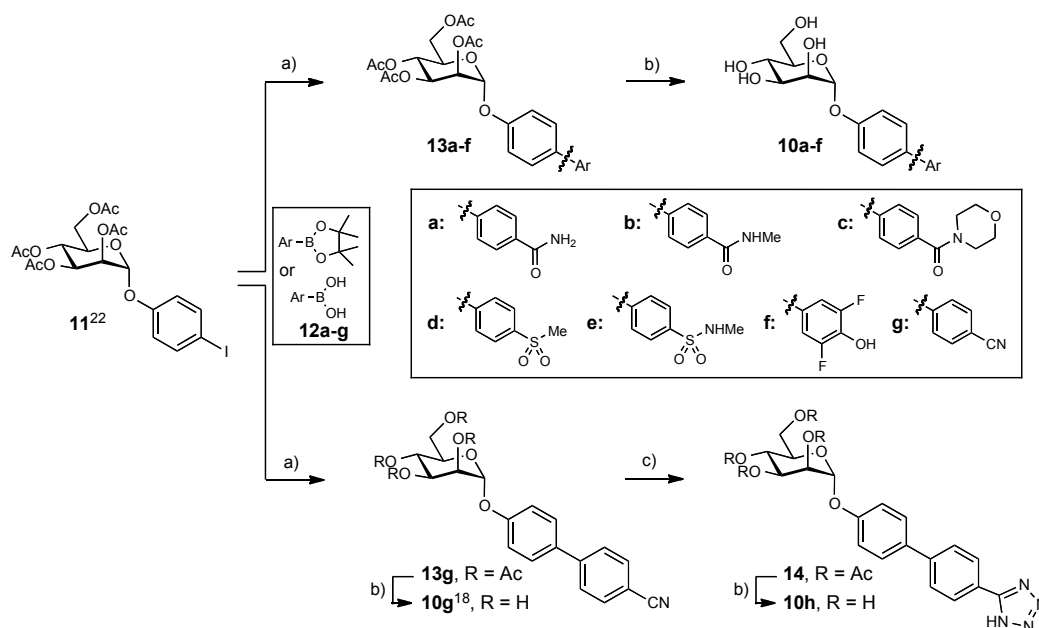


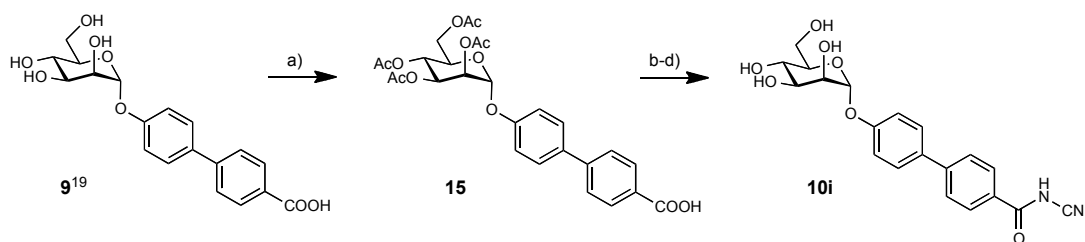
Figure 2. Bioisosteric replacement of the carboxylic acid substituent of biphenyl α -D-mannopyranoside **9**.

Synthesis. Iodide **11** was prepared from peracetylated mannose and 4-iodophenol in the presence of $\text{BF}_3 \cdot \text{Et}_2\text{O}$.²² In a palladium-catalyzed Miyaura-Suzuki coupling⁴⁰ with the boronic acid or boronate derivatives **12a-g**, the biphenyl derivatives **13a-g** were obtained in good to excellent yields. Final deprotection yielded the test compounds **10a-g**. Utilizing microwave-assisted reaction conditions,⁴¹ the conversion of aryl nitrile **13g** to tetrazole **14** proceeded rapidly and with good yield. After deprotection of **14** using Zemplén conditions, the test compound **10h** was obtained (Scheme 1).



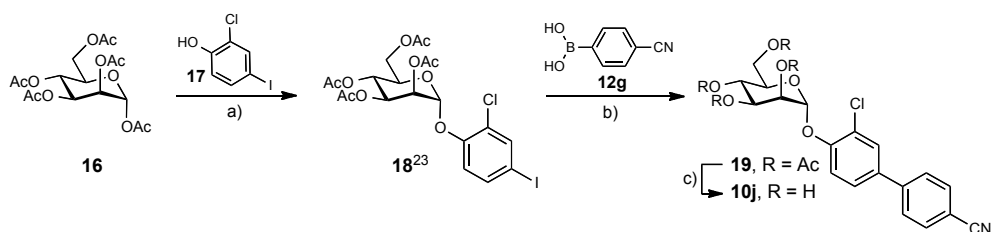
Scheme 1. a) $\text{Pd}(\text{Cl}_2)\text{dppf} \cdot \text{CH}_2\text{Cl}_2$, K_3PO_4 , DMF, 80 °C, 4 h (**13a-g**, 44-99%); b) NaOMe, MeOH, rt, 4 h (**10a-h**, 29-86%); c) TMSN_3 , $\text{Bu}_2\text{Sn}(\text{O})$, DME, 150 °C, μW , 10 min (81%).

The cyanobenzamide derivative **10i** (Scheme 2) was obtained from **9** by peracetylation (\rightarrow **15**) followed by conversion of the carboxylic acid into its acid chloride with 1-chloro-*N,N*,2-trimethyl-1-propenylamine.⁴² Without isolation, the acid chloride was reacted with sodium hydrogen cyanamide in DMF followed by deacetylation under Zemplén conditions to yield the test compound **10i**.



Scheme 2. a) i) Ac_2O , DMAP, pyridine, 0 °C to rt, overnight; ii) satd. NaHCO_3 aq., DCM, rt, 2 h (**15**, 53%); b) 1-chloro-*N,N*,2-trimethyl-1-propenylamine, toluene, 0 °C to rt, 2 h; c) NaH , NH_2CN , DMF, 0 °C to rt, overnight; d) NaOMe , MeOH, rt, 4 h (**10i**, 21% for three steps).

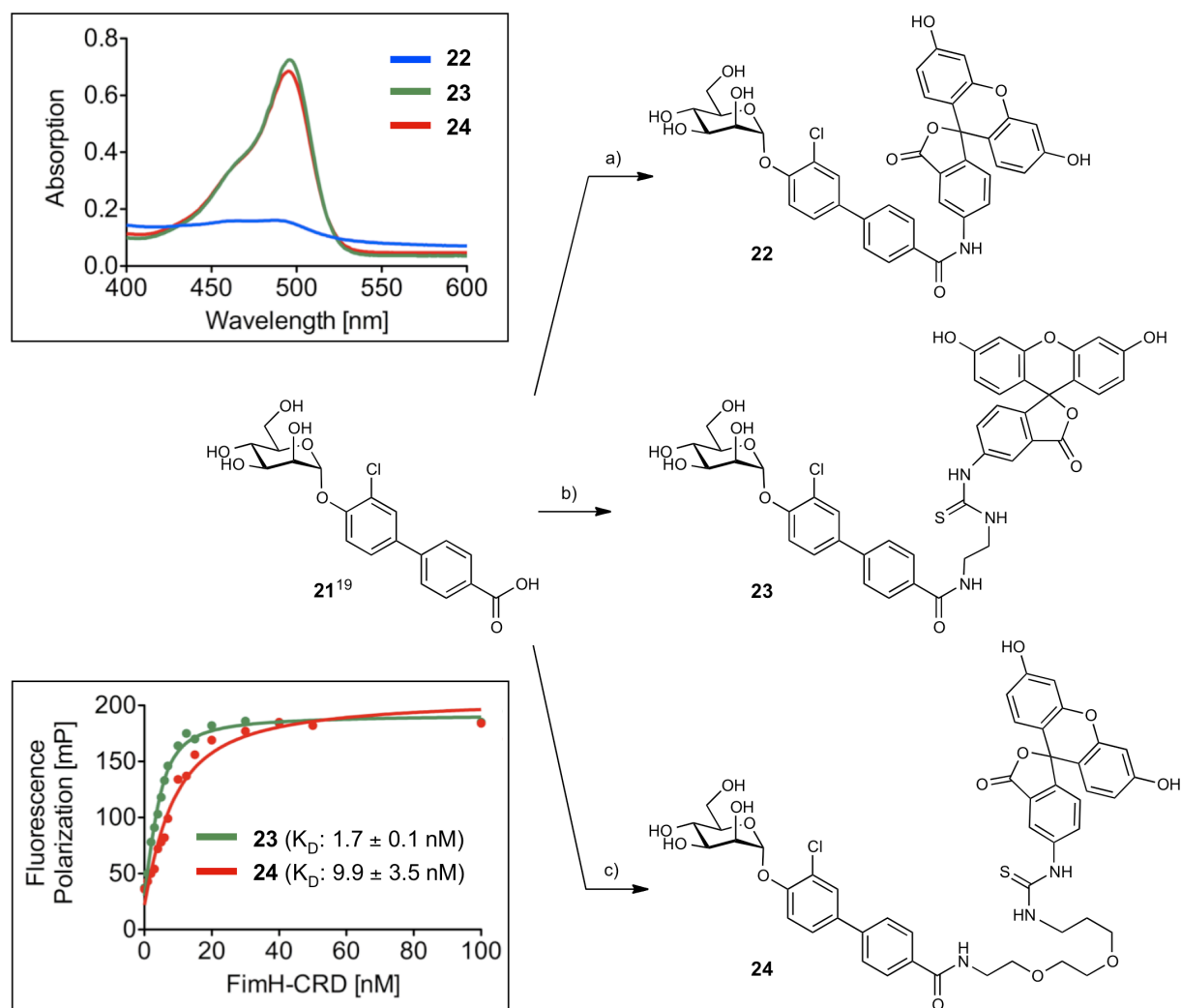
Finally, to further improve the pharmacokinetic properties of mannoside **10g**¹⁸ (see Table 3), a chloride substituent was introduced to the *ortho*-position of the aromatic ring adjacent to the anomeric oxygen. For its synthesis, peracetylated α -D-mannose (**16**) was coupled with 2-chloro-4-iodophenol (**17**) using $\text{BF}_3 \cdot \text{Et}_2\text{O}$ as promotor (\rightarrow **18**, 76%). After the introduction of the second aromatic ring by Miyaura-Suzuki coupling (\rightarrow **19**, 75%), deprotection yielded mannoside **10j** (Scheme 3).



Scheme 3. a) $\text{BF}_3 \cdot \text{Et}_2\text{O}$, CH_2Cl_2 , 40 °C (76%); b) $\text{Pd}(\text{Cl}_2)\text{dppf} \cdot \text{CH}_2\text{Cl}_2$, K_3PO_4 , DMF, 80 °C (75%); c) NaOMe , MeOH, rt, 4 h (48%).

Binding Affinity. The binding affinity of heptyl mannoside **1**, the biphenyl mannosides **3**, **9**, **20**,¹⁸ and the bioisosteres **10a-j** was determined in a competitive fluorescence polarization assay (FP-assay) and with isothermal titration calorimetry (ITC). A protein construct consisting of the CRD with a C-terminal His-tag with a thrombin cleavage site (FimH-CRD-Th-His₆) was used for all experiments.⁴³

Competitive Fluorescence Polarization Assay. For the rapid evaluation of binding affinity, we established a competitive binding assay based on fluorescence polarization (FP). Similar formats have been applied before for the detection of carbohydrate-lectin interactions.^{18,44} In this assay, the antagonist of interest displaces a fluorescently labeled competitor from the binding site, thereby causing a reduction in fluorescence polarization.⁴⁵ To identify the optimal competitor, fluorescein isothiocyanate (FITC) was connected to the FimH ligand **21** by three linkers of different lengths (→ **22-24**, Scheme 4). For optimal sensitivity and signal-to-noise ratio, three main parameters need to be considered: (i) the affinity of the competitor should not be impaired by the fluorescent label, (ii) the conformational flexibility of the label upon binding of the competitor to the CRD should be low and (iii) the fluorescence properties of the label should not be affected by the connected ligand.⁴⁶⁻⁴⁸ A change in fluorescence properties was observed for reporter ligand **22** in which the label was linked to the biphenyl agylcone by an amide bond. The absorption spectrum revealed a lack of the characteristic fluorescein absorption peak at 494 nm (Scheme 4), likely due to an extension of the conjugated system to the biphenyl moiety of the ligand. The elongated saturated spacer groups in competitors **23** and **24** ensured that the expected spectral properties of the dye were retained (Scheme 4).



Scheme 4. a) 1-[(1-(Cyano-2-ethoxy-2-oxoethylideneaminoxy)-dimethylamino-morpholinomethylene)] methanaminium hexafluorophosphate (COMU), NEt_3 , fluoresceinamine, DMF, rt, 7 h (**22**, 19%); b) i. DIC, NHS, *N*-Boc-ethylenediamine, DMF, rt, 12 h; ii. TFA, DCM, rt, 10 min (68% over two steps), iii. fluorescein isothiocyanate (FITC), NEt_3 , DMF, rt, 3 h (**23**, 48%); c) i. DIC, NHS, *N*-Boc-PEG2-NH₂, DMF, rt, 14 h; ii. TFA, DCM, rt, 30 min (62% over two steps); iii. FITC, DMF, rt (**24**, 65%).

For the determination of their binding affinity, fixed concentrations of the reporter ligands **23** and **24** were incubated for 24 h with a linear dilution of the FimH-CRD (0-100 nM). FP was measured using a plate reader, with polarized excitation at 485 nm, and emission at 528 nm measured through appropriately oriented polarizers. Fitting the single-site binding function of Cooper⁴⁹ to the observed FP data resulted for compound **23** in a dissociation constant ($K_D = 1.7$

nM, Figure 3A) similar to that of the unlabeled parent compound **21**,¹⁹ whereas **24** showed a five-fold lower affinity (9.9 nM) (Scheme 4). Therefore, the reporter ligand **23** fulfills all characteristics as an optimal competitor and was used for the FP assay.

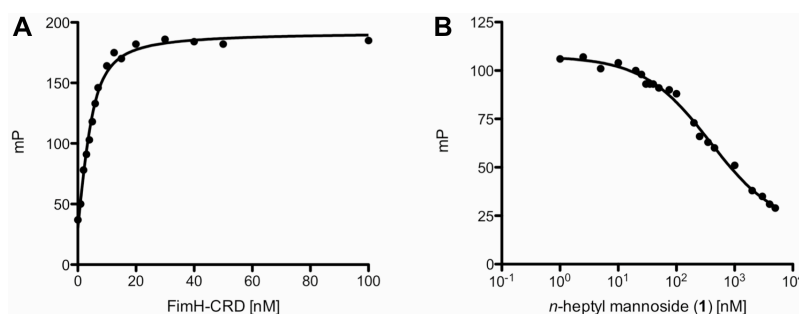


Figure 3. A) Direct binding curve of the labeled competitor **23** obtained by adding a linear dilution of FimH-CRD (0-100 nM) and a constant concentration of competitor **23** (5 nM). The K_D was determined by fitting the experimental data to a single-site binding fit that accounts for ligand depletion. In three FP based direct binding experiments the K_D of competitor **23** was determined to be 1.7 nM. B) Inhibition curve of *n*-heptyl mannoside (**1**) from the competitive FP assay. The IC_{50} value was determined by nonlinear least-squares fitting to a standard 4-parameter equation. A modified Cheng-Prusoff equation⁴⁵ was used to calculate the corresponding K_D value (K_D = 28.3 nM).

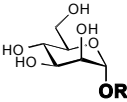
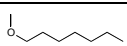
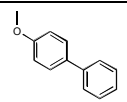
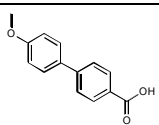
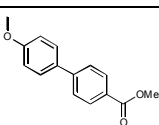
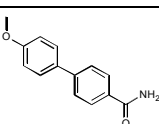
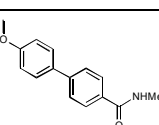
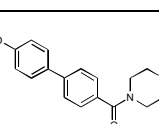
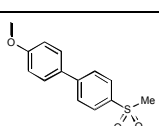
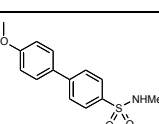
For the test compounds **1**, **3**, **9**, **20**, and **10a-j**, a 24 h incubation time was applied before FP was measured due to the long residence time of FimH antagonists ($t_{1/2} > 3.5$ h, Figure 3B⁵⁰). The 24 h incubation period was empirically determined to be necessary to reach equilibrium between reporter ligand and compound of interest. IC_{50} values were obtained by nonlinear least-squares regression (standard four-parameter dose response curve) and converted to K_D values using a modified Cheng-Prusoff equation.⁴⁵ This equation accounts for the ligand depletion effect in competitive titrations involving high-affinity interaction partners present in similar concentrations. Under these conditions, the free concentration of an interacting species cannot be assumed to equal the total concentration.

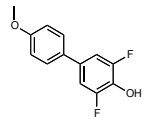
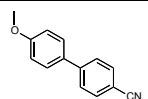
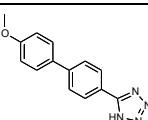
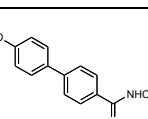
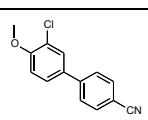
The K_D values determined for the test compounds **1**, **3**, **9**, **20**, and **10a-j** are summarized in Table 1. Against our expectations, the biphenyl mannosides **3** and **9** exhibit similar affinities (Table 1), despite the presence of an electron withdrawing carboxylate substituent in antagonist **9**. According to the crystal structure of FimH co-crystallized with the sulfonamide derivative **10e** (Figure 4A), the outer aromatic ring of the biphenyl aglycone forms π - π interactions with the electron rich Tyr48, which is part of the tyrosine gate of FimH.¹⁵ A reduction of electron density of the aglycone by the electron withdrawing carboxylate was expected to enforce these π - π stacking interactions and lead to improved affinity. However, this beneficial effect might be compensated by an entropic penalty originating from the improved π - π stacking to Tyr48 that might lead to the reduced flexibility of both protein and antagonist. Furthermore, a beneficial enthalpy effect might be partially compensated by an enthalpy penalty originating from the desolvation of the charged carboxylate in **9**⁵¹ (see also experimental part). Although this substituent is solvent exposed, at least a partial desolvation may be necessary upon antagonist binding. To prove this assumption, we replaced the carboxylate by the corresponding methyl ester (\rightarrow **20**)¹⁹ in order to reduce the desolvation penalty and, as predicted by the Hammett constant σ_p ,⁵² to further improve the π - π stacking. Indeed, a six-fold improvement in affinity was achieved. However, since the methyl ester undergoes rapid enzyme-mediated hydrolysis *in vivo*,¹⁹ it will not be available at the place of action in the urinary bladder. The methyl ester was therefore replaced by metabolically stable bioisosteres³⁹ exhibiting comparable electron withdrawing properties⁵² (Table 1, entries 5-13). The most potent derivatives **10d**, **10e** and **10g** showed affinities in the low nanomolar range.

As previously reported,²² a chloro substituent in the *ortho*-position of the aromatic ring adjacent to the anomeric oxygen is favorable for affinity and improves the physicochemical

properties relevant for oral bioavailability. Indeed, the corresponding antagonist **10j** was the most potent compound tested in this study.

Table 1. Affinities (K_D) of FimH antagonists to FimH-CRD-Th-His₆; dissociation constants (K_D) were determined in a competitive fluorescence polarization assay.

Entry	Compd		Affinity K_D [nM]
1	1		28.3 ± 5.0
2	3		15.1 ± 2.2
3	9		17.9 ± 1.5
4	20		3.6 ± 0.9
5	10a		2.8 ± 0.3
6	10b		2.9 ± 0.5
7	10c		3.0 ± 0.1
8	10d		1.7 ± 0.2
9	10e		2.7 ± 0.4

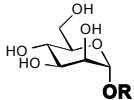
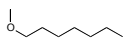
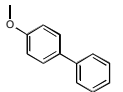
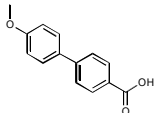
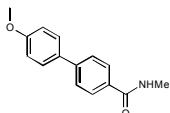
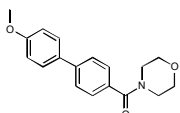
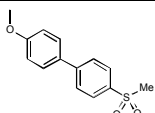
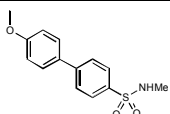
10	10f		3.7 ± 0.2
11	10g		2.0 ± 0.6
12	10h		5.7 ± 0.1
13	10i		8.4 ± 0.3
14	10j		$< 1^a$

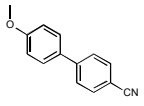
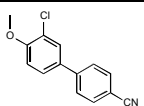
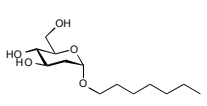
^a) The K_D value of **10j** was approximated to be in the subnanomolar range. The IC_{50} value obtained in the competitive FP assay was equal to the lowest value that can be resolved by the assay, indicating stoichiometric titration of **10j** due to its high affinity. Consequently, its K_D must be below the K_D of competitor **23**.

Isothermal Titration Calorimetry (ITC). To further confirm our hypothesis regarding π - π stacking and desolvation, we performed ITC experiments with the reference compound **1**, the unsubstituted biphenyl mannoside **3**, the carboxylic acid **9**, and the bioisosteres **10b-e**, **g** and **j** (Table 2). ITC allows the simultaneous determination of the stoichiometry (N), the change in enthalpy (ΔH) and the dissociation constant (K_D) for ligand-protein binding.^{53,54} The reliable determination of these three parameters requires well-defined sigmoidal titration curves characterized by the dimensionless Wiseman parameter c ($c = Mt(0) K_D^{-1}$, where $Mt(0)$ is the initial macromolecule concentration).⁵⁵ To be sure that data can be fitted with confidence, the c -value should be between 1 and 1,000 (ideally between 5 and 500),⁵⁶ which could be achieved for the antagonists **3** and **9**. For titrations involving low micromolar $Mt(0)$ and interactions in the low nanomolar or picomolar range, as suggested for the bioisosteres **10b-j**, c -values above

1,000 were expected. Since these conditions lead to steep titration curves that do not allow the determination of the curve slope representing $1/K_D$, we applied an alternative, competitive format referred to as displacement assay.^{57,58} First, FimH-CRD-Th-His₆ was pre-incubated with the low affinity antagonist *n*-heptyl 2-deoxy- α -D-mannopyranoside (**25**, for synthesis see supporting information). The high-affinity bioisosteres of interest were titrated into the protein-ligand complex giving well-defined sigmoidal titration curves.

Table 2. Thermodynamic parameters from ITC for selected FimH-antagonists binding to FimH-CRD-Th-His₆; *n*, stoichiometric correction factor; CI, confidence interval from fitting.

Entry	Compd		K_D ^[a] [nM]	ΔG [kJ/mol]	ΔH ^[a] [kJ/mol]	$-T\Delta S$ [kJ/mol]	<i>n</i>	Type of measurement
1	1 ^[b,c]		28.9 (25.8 – 32.3)	-43.0	-50.3 (-50.2 – -50.7)	7.3	1.00	direct
2	3 ^[b]		17.7 (14.1 – 22.3)	-44.2	-45.0 (-44.5 – -45.6)	0.8	1.07	direct
3	9		15.0 (13.4 – 16.7)	-44.7	-48.7 (-48.4 – -49.0)	4.0	1.05	direct
4	10b		4.3 (3.2 – 5.6)	-47.8	-54.5 (-54.1 – -54.9)	6.7	1.02	competitive vs. 25
5	10c		5.0 (3.8 – 6.6)	-47.4	-54.5 (-54.1 – -54.8)	7.1	0.97	competitive vs. 25
6	10d		3.0 (2.1 – 4.2)	-48.7	-52.3 (-51.5 – -53.1)	3.6	0.99	competitive vs. 25
7	10e		3.5 (2.9 – 4.3)	-48.2	-52.2 (-51.6 – -52.8)	3.9	1.06	competitive vs. 25

8	10g		2.8 (2.3 – 3.3)	-48.8	-58.2 (-57.8 – -58.6)	9.4	1.00	competitive vs. 25
9	10j		1.3 (1.1 – 1.6)	-50.7	-60.9 (-60.4 – -61.4)	10.1	1.01	competitive vs. 25
10	25		9'386 (8'555 – 10'287)	-28.7	-19.5 (-19.1 – -20.0)	-9.1	1.00	direct

^[a] 95% confidence interval from fitting in parentheses; ^[b] Global fit including two direct titration measurements;

^[c] ITC data were previously published with an *n*-value of 0.82.³⁷

The resulting K_D values (Table 2) correspond well with the data obtained from the FP assay (Table 1). A comparison of the thermodynamic fingerprints of antagonists **3** and **9** reveals that the more favorable enthalpic contribution resulting from facilitated π - π stacking leads to a net enthalpy gain ($\Delta\Delta H$: -3.7 kJ/mol). However, an even greater increase in enthalpy is likely countered by the enthalpy costs for desolvation of the electron withdrawing carboxylate.

The gain in enthalpy is in turn compensated by an unfavorable entropy ($-T\Delta\Delta S$: 3.2 kJ/mol) as a result of the reduced flexibility of both the antagonist and the Tyr48 side-chain caused by the improved interaction. This is not entirely outweighed by the beneficial entropy contribution related to the partial desolvation of the carboxylate and the related release of water into the bulk. Added together, the enthalpy and entropy contributions of antagonists **3** and **9** result in similar affinities (K_D : 17.7 and 15.0 nM, respectively).

In contrast, the replacement of the carboxylate group by various neutral bioisosteres (entries 4-7) reduces the enthalpy costs for desolvation (see calculated free energies of desolvation, experimental part) and therefore leads to a markedly improved enthalpy ($\Delta\Delta H$: -3.5 to -5.8

kJ/mol). As a result, an up to fivefold improvement of the K_D values was achieved. Finally, with a cyano substituent (entries 8 & 9), the enthalpy term was further improved ($\Delta\Delta H$: -3.7 kJ/mol) due to a reduced desolvation penalty and improved π - π stacking interactions. However, this beneficial component is again partially compensated by a decrease in entropy. This can be attributed, first, to the loss of flexibility of the tightly bound ligand (Figure 4B) and, second, to the smaller surface area of the cyano substituent compared to amide, sulfonamide and sulfone, which results in a smaller number of water molecules being released to bulk upon binding.

X-ray Crystallography. To determine the binding poses of the bioisosters, we co-crystallized the compounds **10e** and **10j** with the FimH-CRD (Figure 4). Atomic resolution crystal structures were obtained at 1.07 Å (**10e**) and 1.10 Å (**10j**). As observed in previous mannoside co-crystal structures,^{15,18,36} the mannose moiety forms an extensive hydrogen bond network to the well-defined binding site with all of its hydroxyl groups. The biphenyl aglycone is located between the tyrosine gate residues (Tyr48/Tyr137). The π - π stacking of the second aromatic ring of the aglycone to the side chain of Tyr48 contributes most to the interaction energy of the aglycone moiety. Interactions to the Tyr137 side-chain on the other hand are only limited. Whereas a previously published crystal structure of a biphenyl mannoside in complex with FimH-CRD suffers from crystal contacts of binding site residues (Tyr48 side-chain to backbone oxygen of Val27) possibly causing the distortion of the binding site,¹⁸ the binding site of our structures are mostly solvent exposed. This revealed the flexibility of the aglycone in the FimH-CRD/**10e** structure, since the electron density towards the solvent-exposed sulfonamide indicates that there is not one single orientation. Therefore, the aglycone was modeled in two distinct poses. In contrast, in the FimH-CRD/**10j** structure the amino acid side chain of Y48 can be modeled in two distinct rotamers, suggesting flexibility also of the receptor.

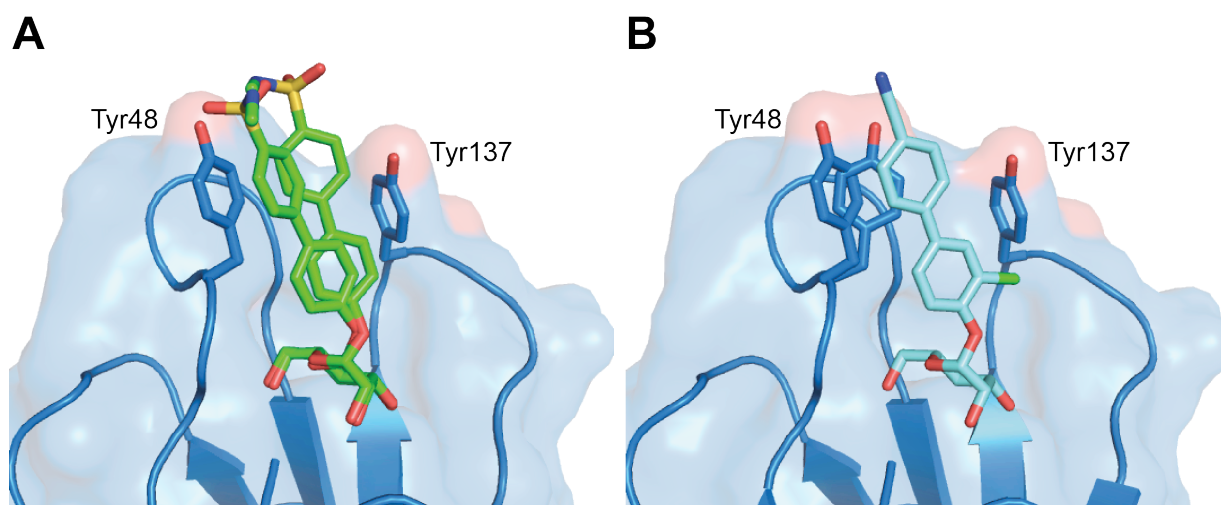


Figure 4. Ligand binding poses determined by X-ray co-crystallization with compounds **10e** resolved to 1.07 Å (**A**) and **10j** resolved to 1.10 Å (**B**). The electron density surrounding the aglycone of **10e** indicates flexibility of the aglycone and was modeled in two poses. Both compounds bind in a similar pose with a well-defined hydrogen network surrounding the mannose moiety and π - π stacking interactions between the second aromatic ring and Tyr48 side-chain (**A**). In contrast, in the FimH-CRD/**10j** structure the amino acid side chain of Y48 can be modeled in two distinct rotamers, suggesting flexibility also of the receptor (**B**).

Physicochemical Properties and *In Vitro* Pharmacokinetics. Intestinal absorption and renal excretion are prerequisites for a successful oral treatment of UTI with FimH antagonists. Furthermore, reabsorption of antagonist from the renal ultrafiltrate is desirable for maintaining the minimal anti-adhesive concentration in the target organ, namely the bladder, over an extended period of time. To estimate the influence of the bioisostere approach on oral bioavailability and the rate of renal excretion, we determined lipophilicity by means of the octanol-water distribution coefficient ($\log D_{7.4}$),⁵⁹ aqueous solubility, and membrane permeability in the artificial membrane permeability assay (PAMPA)⁶⁰ and the colorectal adenocarcinoma (Caco-2) cell monolayer model.⁶¹

Table 3. Physicochemical and *in vitro* pharmacokinetic parameters.

Compd	$pK_a^{a)}$	$\log D_{7.4}^{b)}$	Solubility [$\mu\text{g/mL}$] / $\text{pH}^{c)}$	PAMPA $\log P_e$ [cm/s] / $\text{pH}^{d)}$	Caco-2 P_{app} [10^{-6} cm/s] ^{e)}		PPB f_b [%] ^{f)}	Metabolic stability $t_{1/2}$ [min] ^{g)}
					a→b	b→a		
1	---	1.65	> 3000	-4.89	7.0 ± 0.6	9.4 ± 0.2	81	13
3	---	2.1 ± 0.1	21 ± 1 / 7.4	-4.7 ± 0.1 / 7.4	10.0 ± 0.9	19.0 ± 1.2	93 ± 1	n.d.
20	---	2.14	33.8 / 6.51	-4.7	4.23	n.d.	93	1.0
9	3.88	< -1.5	> 3000 / 6.61	no permeation	n.d.	n.d.	73	> 60
10a	---	0.5 ± 0.1	12 ± 1 / 7.4	-6.8 ± 0.3 / 7.4	0.12 ± 0.01	0.61 ± 0.03	n.d.	n.d.
10b	---	0.8 ± 0.0	122 ± 13 / 7.4	-9.2 ± 1.4 / 7.4	1.10 ± 0.82	0.87 ± 0.15	n.d.	n.d.
10c	---	0.2 ± 0.1	> 250 / 7.4	-7.8 ± 0.3 / 7.4	0.18 ± 0.07	1.30 ± 0.03	48 ± 2	> 60
10d	---	0.4 ± 0.0	246 ± 17 / 7.4	-7.2 ± 0.0 / 7.4	0.36 ± 0.01	1.76 ± 0.12	99 ± 1	> 60
10e	---	0.7 ± 0.1	> 250 / 7.4	-8.6 ± 0.2 / 7.4	0.28 ± 0.23	1.82 ± 0.14	> 99	> 60
10f	6.5	1.1 ± 0.0	> 150 / 3.0 > 150 / 7.4	-7.7 ± 0.8 / 5.0 -8.8 ± 0.1 / 7.4	0.40 ± 0.02	1.90 ± 0.17	n.d.	n.d.
10g	---	1.4 ± 0.0	186 ± 4 / 7.6	-5.7 ± 0.0 / 7.4	2.0 ± 0.1	13.2 ± 2.1	99 ± 0	> 60
10h	3.7	-1.4 ± 0.1	11 ± 0 / 3.0 273 ± 2 / 7.4	-9.3 ± 1.4 / 5.0 -8.8 ± 1.4 / 7.4	0.17 ± 0.00	0.22 ± 0.01	n.d.	n.d.
10i	2.5	-1.1 ± 0.1	> 150 / 3.0 > 150 / 7.4	-6.8 ± 0.2 / 5.0 -7.0 ± 0.1 / 7.4	0.22 ± 0.14	0.29 ± 0.03	n.d.	n.d.
10j	---	2.1 ± 0.0	192 ± 5 / 7.4	-5.2 ± 0.0 / 7.4	2.2 ± 0.4	22.1 ± 1.5	89 ± 1	> 60

a) pK_a values were determined by NMR spectroscopy; b) Octanol-water distribution coefficients ($\log D_{7.4}$) were determined by a miniaturized shake-flask procedure at pH 7.4, values represent the mean \pm SD of sextuplicate measurements;⁵⁹ c) Kinetic solubility was measured in a 96-well format using the μSOL Explorer solubility analyzer at the indicated pH in triplicate; d) P_e = effective permeability: passive permeation through an artificial membrane was determined by the parallel artificial membrane permeation assay (PAMPA), values represent the mean \pm SD of quadruplicate measurements performed at the indicated pH;⁶⁰ e) P_{app} = apparent permeability: permeation through a Caco-2 cell monolayer was assessed in the absorptive (a→b) and secretory (b→a) directions

in triplicate;⁶¹ f) Plasma protein binding (PPB) was determined by equilibrium dialysis in triplicate;⁶² g) Metabolic stability was determined by incubating the compounds (2 μ M) with pooled rat liver microsomes (RLM, 0.5 mg/mL) in presence of NADPH (1 mM, compounds **1**, **9**, **10c-e**, **g**, **j**) or without NADPH (compound **20**);⁶³ n.d. = not determined.

Oral Bioavailability. Oral bioavailability of a compound relies on solubility, permeation through the membranes lining the intestine, and stability against first pass metabolism.^{64,65} As discussed by Lipinski⁶⁶ and Curatolo,⁶⁷ dose and permeability define the minimum aqueous solubility required for oral administration. Thus, a dose of 1 mg/kg of a moderately permeable compound requires a solubility of at least 52 μ g/mL. Whereas sufficient aqueous solubility (> 3000 μ g/mL) was reported for *n*-heptyl α -mannopyranoside (**1**),¹⁹ the unsubstituted biphenyl α -D-mannopyranoside **3** as well as the antagonists bearing a methylcarboxylate, carboxamide, or tetrazole substituent (compounds **20**, **10a** and **10h**) were found to be scarcely soluble.²² As proposed by Ishikawa,⁶⁸ a possible reason is the apolar and planar aglycone. By contrast, the polar carboxylic acid moiety present in antagonist **9** or the substituents in the bioisosteres **10b-j** enhance solubility to 122-273 μ g/mL, a level sufficient for *in vivo* PK studies. For *in vivo* disease studies, however, dosages of up to 10 mg/kg were foreseen (see below), requiring a solubility of 520 μ g/mL.^{66,67} For this reason, surfactant Tween 80 (1%) had to be added.

Furthermore, permeability data derived from PAMPA⁶⁹ and the Caco-2 model⁷⁰ suggest moderate to high permeation of the moderately lipophilic antagonists **1**, **3** and **20** ($\log D_{7.4} > 1.6$) through the intestinal membranes. The bioisosteres **10a-f**, **h**, **i**, although slightly more permeable than the strongly hydrophilic carboxylic acid derivative **9**, show only low values of permeability compared to *n*-heptyl α -D-mannopyranoside (**1**) or the unsubstituted biphenyl mannoside **3**. However, the *para*-cyanobiphenyl derivatives **10g** and **10j** display elevated $\log D_{7.4}$ and effective permeability ($\log P_e$) in the range for successful intestinal absorption.

Regarding both, sufficient aqueous solubility and elevated membrane permeability, the *para*-cyano substituted bioisosteres **10g** and **10j** are thus the most promising candidates for oral absorption. Moreover, combining the bioisosteric replacement with the addition of a chloro substituent in the *ortho*-position of the aromatic ring adjacent to the anomeric oxygen (\rightarrow **10j**)²² resulted in the most advantageous physicochemical profile for oral bioavailability.

Renal Excretion. The rate of renal excretion depends on the rate of glomerular filtration and the propensity to tubular secretion and reabsorption of an antagonist.⁷¹ Only the fraction that is not bound to plasma proteins is expected to enter the glomerular filtrate.⁷² Plasma protein binding (PPB) data indicating the fraction bound (f_b) are listed in Table 2.⁶² The biphenyls **9** and **10c** were identified as moderate binders to plasma proteins ($f_b \leq 65\%$), which suggests a low impact of PPB on antagonist filtration. The f_b values of the antagonists **1**, **3**, **20** and **10j** were between 80 and 93%, whereas the bioisosteres **10d**, **e** and **g** showed particularly high protein binding ($f_b \geq 99\%$) implying slow compound entry into the primary urine. However, the kinetic aspects of PPB, that is, association and dissociation rate constants, remain to be determined to quantify precisely the influence of PPB on filtration.⁷³

Furthermore, $\log D_{7.4}$ was identified as key determinant of tubular reabsorption.⁷⁴⁻⁷⁶ Accordingly, lipophilic compounds are predominantly reabsorbed from the renal filtrate. Given that renal clearance is the major route of elimination, this will result in a slow but steady excretion into the bladder. In contrast, hydrophilic compounds are poorly reabsorbed and thus quickly renally eliminated, which leads to high initial compound levels in the urine but narrows the time range where the minimal anti-adhesive concentration is maintained. Consequently, low $\log D_{7.4}$ as shown for the antagonists **9**, **10h** and **10i** implies low tubular reabsorption and rapid elimination of the filtered molecules by the urine. Otherwise, $\log D_{7.4}$

between 0.2 and 0.7, such as determined for the bioisosteres **10a-e**, suggests increasing propensity to tubular reuptake, whereas $\log D_{7.4} > 1$ as shown for heptyl mannoside **1** and the biphenyl mannosides **3**, **20**, **10g**, **10f** and **10j** is optimal for tubular reabsorption from the glomerular filtrate and thus for slow renal clearance.

Metabolic Stability. Increasing lipophilicity is usually paralleled by increasing susceptibility to metabolism.⁷⁷ Liabilities towards metabolic clearance pathways which prevent the intact antagonist from reaching the target in the bladder were therefore of interest. To assess their propensity to cytochrome P450 (CYP450)-mediated metabolism, heptyl mannoside **1**, the carboxylic acid derivative **9**, and the bioisosteres **10c-e**, **g**, **j** were incubated with rat liver microsomes (RLM, 0.5 mg/mL) in presence of the cofactor β -nicotinamide adenine dinucleotide phosphate (NADPH).⁶³ To confirm the high propensity of the methyl ester present in antagonist **20** to carboxylesterase (CES)-mediated hydrolysis, this antagonist was incubated with RLM only. The profiles of unchanged compound versus time revealed high susceptibility of heptyl mannoside **1** to CYP450-mediated metabolism ($t_{1/2} = 13$ min) and rapid hydrolysis of the ester **20** by the hepatic CES ($t_{1/2} = 1.0$ min). Otherwise, the bioisosteres **10c-e**, **g** & **j** were stable against enzyme-mediated bioconversion ($t_{1/2} > 60$ min) suggesting lower propensity to metabolic, non-renal elimination pathways.

Considering PPB, lipophilicity, and metabolic stability data, we therefore expected (i) a steady release of compounds **10d**, **e**, **g**, **j** into the bladder because of high PPB decelerating glomerular filtration (**10d**, **e**, **g**) and/or high $\log D_{7.4}$ supporting tubular reabsorption (**10g**, **j**), (ii) a fast excretion of antagonists **9** and **10c** via the urine due to low PPB and low $\log D_{7.4}$, and (iii) a rapid clearance of heptyl mannoside **1** from the body by renal and metabolic pathways.

Compounds featuring high propensity to renal excretion as major route of elimination (**10c**, **10e** and **10j**) were selected for *in vivo* PK studies in a mouse model.

Pharmacokinetic Studies in C3H/HeN Mice. This first part of our study explored the predicted effects of lipophilicity, PPB, and metabolic stability on antagonist disposition and elimination upon a single dose iv application (50 mg/kg) of compounds **10c** and **10e**. The PK parameters of these applications and those of the previously published carboxylate **9** are summarized in Table 4. The Table also contains the results of the iv administration of compound **10j** (0.625 mg/kg).

Table 4. Pharmacokinetic parameters determined after a single iv application of compounds **9**, **10c**, **10e** and **10j** in female C3H/HeN mice. Values were calculated using PKSolver.⁷⁸ C_0 , initial concentration; V_z , volume of distribution in terminal phase; AUC, Area under the curve; CL_{tot} , total clearance; C_{max} , maximal concentration.

Compd	Plasma						Urine
	C_0 (μg/mL)	Dose (mg/kg)	V_z (mL)	$t_{1/2}$ (h)	AUC _{0-inf} (μg x h/mL)	CL_{tot} (mL/h)	C_{max} (μg/mL)
9	40	50	25.2	0.33	23.5	53.1	300
10c	109.7	50	28.3	0.4	25.3	49.4	4611
10e	151.6	50	19.5	1.9	175.1	7.1	387
10j	0.36	0.625	52.8	0.17	0.07	218	10

In contrast to the fast plasma clearance of antagonists **9** and **10c** (Figure 5A), the methyl sulfonamide bioisostere **10e** attained higher initial concentration in plasma (C_0) and lower total clearance (CL_{tot}). Therefore, it could be detected until 6 h post application, resulting in markedly higher plasma AUC. The observed high C_0 of compound **10e** may be attributed to a small volume of distribution (V_z) resulting from the high PPB ($f_b \geq 99\%$).⁷² In urine (Figure 5B), the carboxylic acid **9** and the morpholinomethanone **10c** displayed high levels immediately following administration and a rapid concentration decrease within the first two

hours, reflecting the rapid elimination from plasma. Fast renal excretion as major route of elimination can be rationalized by the physicochemical properties of the antagonists **9** and **10c**, that is, moderate PPB and $\log D_{7.4}$, as well as high metabolic stability. Otherwise, the methyl sulfonamide bioisostere **10e** showed sustained compound levels in urine over a period of 2 h and subsequent slow decrease until 6 h post administration. This sustained renal excretion is a result of the interplay of the antagonist's elevated PPB and $\log D_{7.4}$.

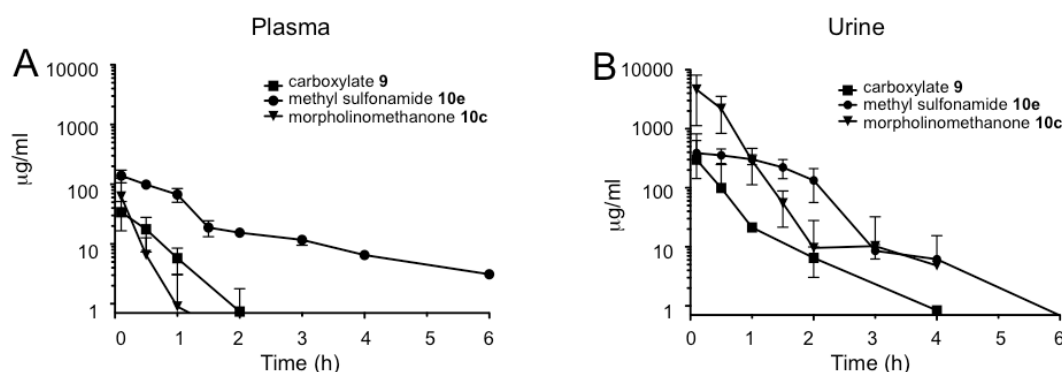


Figure 5. Antagonist concentrations in (A) plasma and (B) urine after a single iv application of **9**, **10c**, and **10e** (50 mg/kg).

In a second study, the *para*-cyano bioisostere **10j**, characterized by a high oral absorption potential, was administered as a single dose iv (0.625 mg/kg) and po (1.25 mg/kg). The plasma concentration curve upon iv dosing displays a steep decline within the first hour post application, while the po curve shows a prolonged period where absorption and elimination are in equilibrium (Figure 6A). The urine concentration profiles (Figure 6B) parallel the plasma curves obtained by the two modes of application, i.e. high plasma clearance upon iv bolus injection led to high initial antagonist levels in urine and a rapid concentration decline. By contrast, sustained plasma concentrations upon po administration resulted in prolonged urine levels.

As a result, urine concentrations exceed the minimum level required for the anti-adhesive effect as estimated from the *in vitro* cell infection model⁷⁹ (minimal anti-adhesion concentration²³, MAC₉₀, 0.094 µg/mL) for more than eight hours upon oral single-dose administration (Figure 6B).

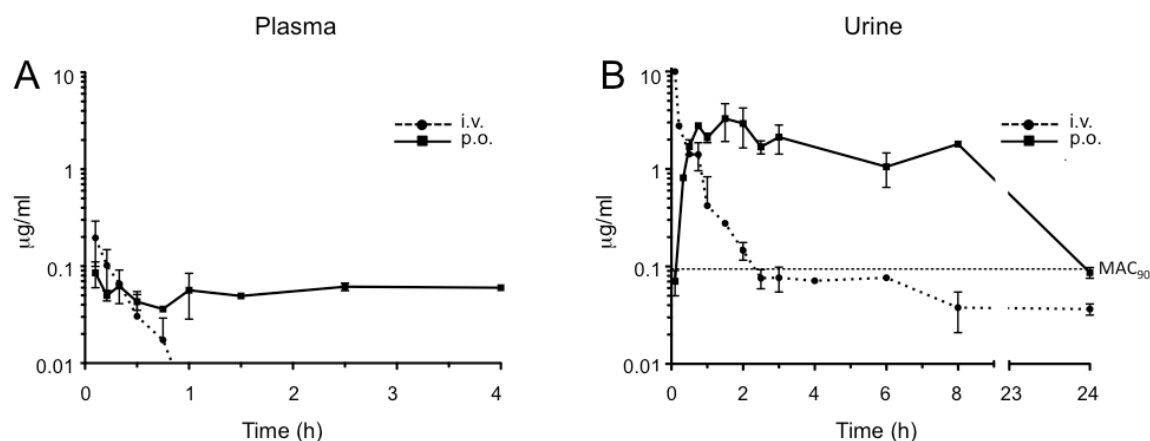


Figure 6. Antagonist concentrations in (A) plasma and (B) urine after a single iv and po application of compound **10j** (iv: 0.625 mg/kg; po: 1.25 mg/kg). MAC₉₀, minimal anti-adhesive concentration to inhibit 90% adhesion (0.094 µg/mL).

Infection study in C3H/HeN Mice. Based on the promising PK profile of **10j**, six mice were inoculated with UTI89 following an oral application of **10j** (1.25 mg/kg) 40 min prior to infection. Three hours after inoculation, the animals were sacrificed and bladder and kidneys were removed. Organs were homogenized and analyzed for bacterial counts. The effect of the FimH antagonist was compared to a 8 mg/kg dose of ciprofloxacin (CIP) subcutaneously (s.c.) applied, used as standard antibiotic therapy against UTI89.⁸⁰ The median reduction in bacterial counts of the mice treated with **10j** and CIP compared to the control group are displayed in Figure 7.

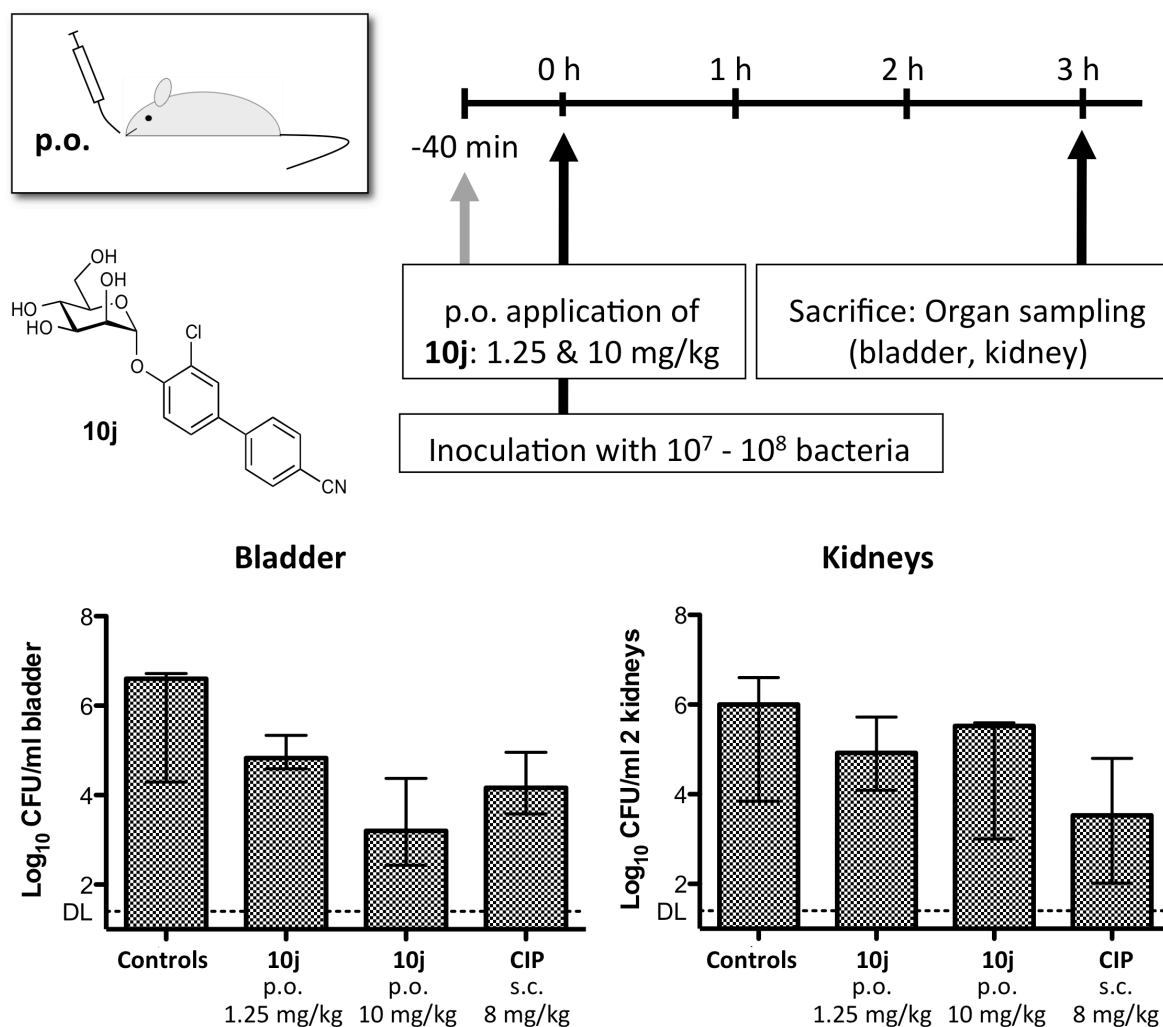


Figure 7. Treatment efficacy in the UTI mouse model 3 h after infection. The bars depict the median bacterial load with the interquartile range in the different study groups. Shown are the results of the untreated control group and the intervention groups with applications of either 1.25 mg/kg **10j**, 10 mg/kg **10j** p.o. or 8 mg/kg CIP s.c. (representing the murine dose equivalent to a human standard dose).⁸¹ **10j** in a dose of 10 mg/kg results in a median reduction of over 3 log₁₀ CFU/ml of bacterial counts in the bladder.

The median value in the untreated control group showed bacterial counts of 6.6 log₁₀ colony forming units (CFU) in the bladder and 6 log₁₀ CFU in the kidneys. After oral application of 1.25 mg/kg of **10j**, bacterial loads in the bladder decreased by 1.7 log₁₀ CFU and 1.08 log₁₀ CFU in the kidneys. The lower reduction in the kidneys is most likely due to the differing adhesion mechanisms between bladder and kidneys (type 1 pili vs P-pili), which is not targeted by **10j**.⁸² With CIP (8 mg/kg s.c.) a substantial reduction in both, bladder and kidneys (median

reductions of 2.44 log₁₀ and 2.47 log₁₀, respectively) was observed. Despite the low oral dose of **10j** (1.25 mg/kg), the approximately 100-fold reduction of CFU in the bladder promised an even higher effect upon dose increase to 10 mg/kg. However, solubility of **10j** for this increased dose is too low (192 µg/mL), therefore we used surfactants as solubilizer (Tween 80, 1%) to enhance solubility. When 10 mg/kg were applied, bacterial loads in the bladder decreased by 3.4 log₁₀ CFU/mL, exceeding the effect of CIP, but only 0.5 log₁₀ CFU in the kidneys. Based on this large reduction after an oral application of 10 mg/kg, **10j** has a large potential for further development.

SUMMARY AND CONCLUSION

Recently, numerous monovalent alkyl and aryl α-D-mannopyranosides have been described as potent FimH antagonists. However, most of them suffer from insufficient pharmacokinetic properties, i.e. modest bioavailability and short duration of the therapeutic effect in the bladder, their site of action. As a consequence, high doses at short intervals are required to achieve anti-adhesive effects over an extended period of time. Therefore, the goal of the present study was an appropriate optimization of the pharmacokinetic profile of biphenyl α-D-mannopyranosides while keeping their high affinity to the CRD of FimH. The starting point was the biphenyl-carboxylate **9** where the critical carboxylate was replaced by bioisosteres.^{39,83} With a series of bioisosteres a three- to fivefold improvement of affinity was achieved compared to **9**. Although binding necessitates only partial desolvation of the carboxylate and its bioisosteric replacements, a reduction of the enthalpy penalty for desolvation⁵¹ was identified as the source of the improved affinity exhibited by the bioisosteres. Thermodynamic evaluation of antagonists **10b-e** revealed almost identical enthalpy contribution to binding. However, for antagonists with the *para*-cyano substituent (**10g** & **10j**) an enhancement of up to

-8.7 kJ/mol was observed, indicating a reduced desolvation penalty and an improved stacking as derived from the crystal structure of **10j** co-crystallized with the CRD of FimH (Figure 4B). On the other hand, higher affinity originating from a reduction of conformational flexibility of ligand and protein resulted in a concomitant entropy penalty of up to 6.5 kJ/mol.

In addition to the improved pharmacodynamics, the relevant pharmacokinetic parameters (solubility, permeability, renal excretion) could also be substantially improved. With 3'-chloro-4'-(α -D-mannopyranosyloxy)-biphenyl-4-carbonitrile (**10j**), a FimH antagonist with an optimal *in vitro* PK/PD profile was identified. The *para*-cyano substituent conferred lipophilicity and high binding to plasma proteins, which slowed down the rate of renal excretion. Despite higher lipophilicity, antagonist **10j** was insusceptible to CYP450-mediated metabolism, and therefore predominantly eliminated via the renal pathway. *In vivo* experiments confirmed the excellent PK-profile of **10j** with steady renal excretion for more than 8 h after oral application, suggesting a long-lasting anti-adhesive effect. Finally, orally applied **10j** (10 mg/kg) was effective in a mouse model of UTI by reducing the bacterial load in the bladder by over 1000-fold. Improvement of solubility, enabling the application of higher dosages of **10j**, will possibly lead to an even higher bacterial reduction, which will be the subject of future investigations.

EXPERIMENTAL SECTION

Synthesis

The synthesis of compounds **10a-d**, **10f**, **10g**, **10i**, **13a-d**, **13f**, **13g**, **15**, **18** and **25**, including compound characterization data, can be found in the Supporting Information.

General methods. NMR spectra were recorded on a Bruker Avance DMX-500 (500.1 MHz) spectrometer. Assignment of ^1H and ^{13}C NMR spectra was achieved using 2D methods (COSY, HSQC, HMBC). Chemical shifts are expressed in ppm using residual CHCl_3 , CHD_2OD or HDO as references. Optical rotations were measured using Perkin-Elmer Polarimeter 341. Electron spray ionization mass spectra (ESI-MS) were obtained on a Waters micromass ZQ. The LC/HRMS analysis were carried out using a Agilent 1100 LC equipped with a photodiode array detector and a Micromass QTOF I equipped with a 4 GHz digital-time converter. Microwave-assisted reactions were carried out with a CEM Discover and Explorer. Reactions were monitored by TLC using glass plates coated with silica gel 60 F_{254} (Merck) and visualized by using UV light and/or by charring with a molybdate solution (a 0.02 M solution of ammonium cerium sulfate dihydrate and ammonium molybdate tetrahydrate in aqueous 10% H_2SO_4). MPLC separations were carried out on a CombiFlash Companion or Rf (Teledyne Isco) equipped with RediSep normal-phase or RP-18 reversed-phase flash columns. LC-MS separations were done on a Waters system equipped with sample manager 2767, pump 2525, PDA 2525 and micromass ZQ. All compounds used for biological assays are at least of 95% purity based on HPLC analytical results. Commercially available reagents were purchased from Fluka, Aldrich, Alfa Aesar or abcr GmbH & Co. KG (Germany). Solvents were purchased from Sigma-Aldrich or Acros and were dried prior to use where indicated. Methanol (MeOH) was dried by refluxing with sodium methoxide and distilled immediately

before use. Dimethoxyethane (DME) was dried by filtration over Al₂O₃ (Fluka, type 5016 A basic).

4'-(2,3,4,6-Tetra-*O*-acetyl- α -D-mannopyranosyloxy)-*N*-methyl-biphenyl-4-sulfonamide

(13e). A Schlenk tube was charged with aryl iodide **11**²² (116 mg, 0.21 mmol), 4-(*N*-methylsulfamoyl)-phenylboronic acid (**12e**, 50 mg, 0.23 mmol), Pd(dppf)Cl₂·CH₂Cl₂ (5 mg, 0.006 mmol), K₃PO₄ (67 mg, 0.32 mmol) and a stirring bar. The tube was closed with a rubber septum and was evacuated and flushed with argon. This procedure was repeated once, and then anhydrous DMF (1 mL) was added under a stream of argon. The mixture was degassed in an ultrasonic bath and flushed with argon for 5 min, and then stirred at 80 °C overnight. The reaction mixture was cooled to rt, diluted with EtOAc (50 mL), and washed with water (50 mL) and brine (50 mL). The organic layer was dried over Na₂SO₄ and concentrated *in vacuo*. The residue was purified by MPLC on silica gel (petroleum ether/EtOAc) to afford **13e** (105 mg, 84%) as a white solid. $[\alpha]_D^{20} +56.4$ (*c* 0.50, MeOH); ¹H NMR (500 MHz, CDCl₃): δ = 7.92-7.90 (m, 2H, Ar-H), 7.70-7.68 (m, 2H, Ar-H), 7.57-7.55 (m, 2H, Ar-H), 7.21-7.19 (m, 2H, Ar-H), 5.60-5.57 (m, 2H, H-1, H-3), 5.48 (dd, *J* = 1.8, 3.4 Hz, 1H, H-2), 5.40 (t, *J* = 10.0 Hz, 1H, H-4), 4.38 (dd, *J* = 5.4, 10.8 Hz, 1H, NH), 4.30 (dd, *J* = 4.9, 12.3 Hz, 1H, H-6a), 4.13-4.08 (m, 2H, H-5, H-6b), 2.72 (d, *J* = 5.4 Hz, 3H, NCH₃), 2.22, 2.07, 2.05, 2.04 (4 s, 12H, 4 COCH₃); ¹³C NMR (126 MHz, CDCl₃): δ = 170.55, 170.06, 170.03, 169.75 (4 CO), 155.97, 144.81, 137.16, 134.09, 128.62, 127.85, 127.39, 117.01 (Ar-C), 95.78 (C-1), 69.34 (C-5), 69.31 (C-2), 68.81 (C-3), 65.86 (C-4), 62.07 (C-6), 29.44 (NHCH₃), 20.92, 20.74, 20.72 (4C, 4 COCH₃); ESI-MS: *m/z*: Calcd for C₂₇H₃₁NNaO₁₂S [M+Na]⁺: 616.1, found: 616.1.

4'-(α -D-Mannopyranosyloxy)-*N*-methyl-biphenyl-4-sulfonamide (10e). To a solution of **13e** (40 mg, 0.07 mmol) in dry MeOH (5 mL) was added freshly prepared 1 M NaOMe/MeOH

(0.1 eq) under argon. The mixture was stirred at rt until the reaction was complete (monitored by TLC), then neutralized with Amberlyst-15 (H^+) ion-exchange resin, filtered and concentrated *in vacuo*. The residue was purified by MPLC on silica gel (DCM/MeOH, 10:1-7:1) to afford **10e** (22 mg, 76%) as white solid. $[\alpha]_{\text{D}}^{20} +105.7$ (*c* 0.30, MeOH); ^1H NMR (500 MHz, CD_3OD): δ = 7.90-7.88 (m, 2H, Ar-H), 7.80-7.79 (m, 2H, Ar-H), 7.66-7.64 (m, 2H, Ar-H), 7.26-7.25 (m, 2H, Ar-H), 5.58 (d, J = 1.7 Hz, 1H, H-1), 4.06 (dd, J = 1.8, 3.3 Hz, 1H, H-2), 3.96 (dd, J = 3.4, 9.5 Hz, 1H, H-3), 3.79-3.74 (m, 3H, H-4, H-6a, H-6b), 3.63 (ddd, J = 2.5, 5.2, 9.7 Hz, 1H, H-5), 2.57 (s, 3H, NHCH_3); ^{13}C -NMR (126 MHz, CD_3OD): δ = 158.34, 146.13, 138.67, 134.55, 129.53, 128.82, 128.21, 118.29 (Ar-C), 100.09 (C-1), 75.53 (C-5), 72.42 (C-3), 71.96 (C-2), 68.32 (C-4), 62.68 (C-6), 29.31 (NHCH_3); HRMS: *m/z*: Calcd for $\text{C}_{19}\text{H}_{23}\text{NNaO}_8\text{S}$ $[\text{M}+\text{Na}]^+$: 448.1037, found: 448.1038.

5-(4'-(2,3,4,6-Tetra-*O*-acetyl- α -D-mannopyranosyloxy)-biphenyl-4-yl)-1*H*-tetrazole (14).

A Schlenk tube was charged with **13g** (30 mg, 0.06 mmol), trimethylsilyl azide (16 μL , 0.12 mmol), dibutyltin oxide (2 mg, 0.006 mmol), DME (1 mL) and a stirring bar. The mixture was heated to 150 $^\circ\text{C}$ for 10 min by microwave irradiation. The reaction mixture was cooled to rt, and then concentrated *in vacuo*. The residue was purified by MPLC on silica gel (DCM/MeOH, 9:1-8:1) to afford **14** (26 mg, 81%) as colorless oil. $[\alpha]_{\text{D}}^{20} +56.1$ (*c* 0.3, MeOH); ^1H NMR (500 MHz, CDCl_3): δ = 8.25-8.15 (m, 2H, Ar-H), 7.75-7.65 (m, 2H, Ar-H), 7.60-7.55 (m, 2H, Ar-H), 7.20-7.17 (m, 2H, Ar-H), 5.64-5.55 (m, 2H, H-1, H-3), 5.49 (dd, J = 1.7, 3.3 Hz, 1H, H-2), 5.40 (t, J = 10.1 Hz, 1H, H-4), 4.31 (dd, J = 5.3, 12.4 Hz, 1H, H-6a), 4.17-4.06 (m, 2H, H-5, H-6b), 2.22, 2.07, 2.06, 2.05 (4 s, 12H, 4 COCH_3); ^{13}C NMR (126 MHz, CDCl_3): δ = 170.67, 170.14, 170.11, 169.81 (4 CO), 155.61, 128.36, 127.84, 127.49, 116.93 (Ar-C), 95.78 (C-1), 69.36 (C-5), 69.26 (C-2), 68.90 (C-3), 65.89 (C-4), 62.12 (C-6),

20.92, 20.76, 20.73 (4 COCH₃); ESI-MS: m/z : Calcd for C₂₇H₂₈N₄NaO₁₀ [M+Na]⁺: 591.2, found: 591.1.

5-(4'-(α -D-Mannopyranosyloxy)-biphenyl-4-yl)-1H-tetrazole (10h). Prepared according to the procedure described for **10e** from **14** (26 mg, 0.03 mmol). Yield: 18 mg (quant.) as a white solid. $[\alpha]_D^{20} +112.1$ (c 0.1, MeOH/H₂O, 2:1); ¹H NMR (500 MHz, CD₃OD): δ = 7.98-7.96 (m, 2H, Ar-H), 7.72-7.71 (m, 2H, Ar-H), 7.58-7.54 (m, 2H, Ar-H), 7.16-7.13 (m, 2H, Ar-H), 5.46 (d, J = 1.7 Hz, 1H, H-1), 3.94 (dd, J = 1.9, 3.5 Hz, 1H, H-2), 3.83 (dd, J = 3.4, 9.5 Hz, 1H, H-3), 3.68-3.61 (m, 3H, H-4, H-6a, H-6b), 3.52 (ddd, J = 2.5, 5.4, 9.7 Hz, 1H, H-5); ¹³C NMR (126 MHz, CD₃OD): δ = 158.19, 145.07, 134.97, 129.29, 128.74, 128.55, 118.26 (Ar-C), 100.13 (C-1), 75.52 (C-5), 72.42 (C-3), 71.98 (C-2), 68.33 (C-4), 62.69 (C-6); HRMS: m/z : Calcd for C₁₉H₂₁N₄O₆ [M+H]⁺: 401.1456, found: 401.1450.

4'-(2,3,4,6-Tetra-*O*-acetyl- α -D-mannopyranosyloxy)-3'-chloro-biphenyl-4-carbonitrile (19). Prepared according to the procedure described for **13e** from aryl iodide **18**²³ (79 mg, 0.135 mmol), **12g** (22 mg, 0.15 mmol), Pd(dppf)Cl₂·CH₂Cl₂ (3.3 mg, 4 μ mol) and K₃PO₄ (57 mg, 0.27 mmol). Yield: 57 mg (75%) as a white solid. $[\alpha]_D^{20} +77.7$ (c 0.5, CHCl₃); ¹H NMR (500 MHz, CDCl₃): δ = 7.72 (d, J = 8.3 Hz, 2H, Ar-H), 7.63 (m, 3H, Ar-H), 7.43 (dd, J = 2.2, 8.6 Hz, 1H, Ar-H), 7.27 (d, J = 8.6 Hz, 1H, Ar-H), 5.64-5.59 (m, 2H, H-1, H-2), 5.54 (dd, J = 1.9, 3.2 Hz, 1H, H-3), 5.41 (t, J = 10.1 Hz, 1H, H-4), 4.28 (dd, J = 5.2, 12.3 Hz, 1H, H-6a), 4.17 (ddd, J = 2.1, 5.1, 10.0 Hz, 1H, H-5), 4.10 (dd, J = 2.2, 12.3 Hz, 1H, H-6b), 2.21 (s, 3H, COCH₃), 2.12- 2.00 (m, 9H, 3 COCH₃); ¹³C NMR (126 MHz, CDCl₃): δ = 170.54, 170.08, 169.90, 169.84, (4C, CO) 151.67, 143.61, 135.29, 132.87, 129.41, 127.53, 126.60, 125.20, 118.79, 117.36, 111.47 (Ar-C, CN), 96.72 (C-1), 70.00 (C-5), 69.39 (C-3), 68.82 (C-2), 65.86

(C-4), 62.16 (C-6), 20.98, 20.81, 20.79, 20.78 (4 COCH₃); ESI-MS: m/z : Calcd for C₂₇H₂₆ClNNaO₁₀ [M+Na]⁺: 582.1, found: 582.1.

3'-Chloro-4'-(α -D-mannopyranosyloxy)-biphenyl-4-carbonitrile (10j). Prepared according to the procedure described for **10e** from **19** (36 mg, 0.06 mmol). Yield: 12 mg (48%) as a white solid. $[\alpha]_D^{20}$ +109.4 (c 0.23, MeOH); ¹H NMR (500 MHz, CD₃OD): δ = 7.80-7.72 (m, 5H, Ar-H), 7.59 (dd, J = 2.2, 8.6 Hz, 1H, Ar-H), 7.48 (d, J = 8.7 Hz, 1H, Ar-H), 5.62 (d, J = 1.4 Hz, 1H, H-1), 4.12 (dd, J = 1.8, 3.3 Hz, 1H, H-2), 4.00 (dd, J = 3.4, 9.5 Hz, 1H, H-3), 3.83-3.68 (m, 3H, H-4, H-6a, H-6b), 3.63 (ddd, J = 2.3, 5.4, 9.6 Hz, 1H, H-5); ¹³C NMR (126 MHz, CD₃OD): δ = 153.65, 145.15, 135.42, 133.86, 129.82, 128.53, 127.87, 125.47, 119.70, 118.59 (Ar-C), 111.97 (CN), 100.66 (C-1), 76.05 (C-5), 72.39 (C-3), 71.80 (C-2), 68.20 (C-4), 62.65 (C-6); IR (KBr), ν = 3400 (OH), 2227 (C \equiv N), 1606, 1487 (Ar-C=C) cm⁻¹; HRMS: m/z : Calcd for C₁₉H₁₈ClNNaO₆ [M+Na]⁺: 414.0715, found: 414.0721.

3'-Chloro-N-(3',6'-dihydroxy-3-oxo-3H-spiro[isobenzofuran-1,9'-xanthen]-5-yl)-4'-(α -D-mannopyranosyloxy)-biphenyl-4-carboxamide (22). Compound **21** (10.0 mg, 0.024 mmol), fluoresceinamine isomer I (12.7 mg, 0.037 mmol) and COMU (20.9 mg, 0.049 mmol) were dissolved in dry DMF (1 mL), then NEt₃ (10 μ L, 0.073 mmol) was added and the mixture was stirred at rt for 7 h. 1 N HCl in DMF was added until acid reaction on pH paper and the mixture was concentrated. The residue was dissolved in DCM/MeOH (3:1) and loaded onto a silica gel column. The complex mixture of compounds was only partially resolved. The fractions containing the product were collected, concentrated and purified by preparative HPLC (gradient H₂O/MeCN, +0.2% HCO₂H), to afford compound **22** (5 mg, 19%). $[\alpha]_D^{20}$ +21.1 (c 0.10, MeOH); ¹H NMR (500 MHz, CD₃OD): δ = 8.26 (d, J = 8.4 Hz, 2H, Ar-H), 7.88-7.74 (m, 3H, Ar-H), 7.66 (dd, J = 2.2, 8.6 Hz, 1H, Ar-H), 7.51 (d, J = 8.7 Hz, 1H, Ar-H),

7.29 (dd, $J = 1.9, 5.3$ Hz, 2H, Ar-H), 7.19 (dd, $J = 2.1, 8.3$ Hz, 1H, Ar-H), 7.08-6.99 (m, 2H, Ar-H), 6.95 (d, $J = 8.7$ Hz, 1H, Ar-H), 6.72 (dd, $J = 5.5, 10.6$ Hz, 2H, Ar-H), 6.61 (dd, $J = 2.3, 8.7$ Hz, 1H, Ar-H), 5.65 (s, 1H, H-1), 4.15 (dd, $J = 1.8, 3.2$ Hz, H-2), 4.03 (dd, $J = 3.4, 9.5$ Hz, H-3), 3.87-3.72 (m, 3H, H-4, H-6a, H-6b), 3.65 (m, 1H, H-5); ^{13}C NMR (126 MHz, CD_3OD): $\delta = 137.50, 136.01, 131.90, 130.24, 130.20, 129.87, 129.24, 128.03, 127.91, 125.79, 125.46, 124.73, 118.99, 118.76, 118.65$ (Ar-C), 100.73 (C-1), 76.06 (C-5), 72.42 (C-3), 71.85 (C-2), 68.24 (C-4), 62.69 (C-2); ESI-MS: m/z : Calcd for $\text{C}_{39}\text{H}_{31}\text{ClINO}_{12}$ $[\text{M}+\text{H}]^+$: 740.2, found: 740.2.

3'-Chloro-*N*-(2-(3-(3',6'-dihydroxy-3-oxo-3*H*-spiro[isobenzofuran-1,9'-xanthen]-5-yl)-thioureido)ethyl)-4'-(α -D-mannopyranosyloxy)-biphenyl-4-carboxamide (23). To a stirred solution of compound **21** (25 mg, 0.061 mmol) in dry DMF (1 mL), NHS (21 mg, 0.183 mmol) was added, followed by DIC (9.2 mg, 0.073 mmol). The mixture was stirred at rt for 2 h, then *N*-Boc-ethylendiamine (10.7 mg, 0.067 mmol) was added and the reaction was stirred for 10 h. It was then cooled down to 0 °C, diluted with water and concentrated. Chromatography on silica gel (DCM/MeOH) yielded 23 mg (0.042 mmol, 68%) of *tert*-butyl (3'-chloro-4'-(α -D-mannopyranosyloxy)-biphenyl-4-yl-carboxamido)ethyl)carbamate. This product was dissolved in DCM (3 mL) and TFA (1 mL) was added. The solid dissolved during addition of TFA. After 10 min the reaction was complete. The mixture was evaporated and excess TFA was removed in high vacuum. The intermediate *N*-(2-aminoethyl)-3'-chloro-4'-(α -D-mannopyranosyloxy)-biphenyl-4-carboxamide TFA salt (23 mg, 0.042 mmol, quant.) was used directly in the next step. It was dissolved in dry DMF (0.5 mL) and NEt_3 (12.8 mg, 0.127 mmol) was added. The mixture was cooled to 0 °C, then FITC (14.8 mg, 0.038 mmol) was added and the mixture was stirred for 3 h in the dark. The mixture was then co-evaporated with water, taken up in MeOH/10% aq. acetic acid and evaporated. Chromatography on silica gel (DCM/MeOH) yielded compound **23**, contaminated with triethylammonium acetate. The

compound was then re-dissolved in MeOH, and 0.5 N HCl in MeOH was added. The mixture was evaporated and chromatographed on silica gel, to yield pure **23** (15 mg, 47%). $[\alpha]_D^{20} +12.1$ (c 0.30, MeOH); ^1H NMR (500 MHz, CD_3OD): δ = 8.12 (s, 1H), 7.92 (d, J = 8.3 Hz, 2H, Ar-H), 7.70 (dd, J = 5.0, 13.1 Hz, 2H, Ar-H), 7.64 (d, J = 8.3 Hz, 2H, Ar-H), 7.54 (dd, J = 2.2, 8.6 Hz, 1H, Ar-H), 7.46 (d, J = 8.7 Hz, 1H, Ar-H), 7.09 (d, J = 8.2 Hz, 1H, Ar-H), 6.74 (s, 2H), 6.69 (d, J = 1.4 Hz, 2H, Ar-H), 6.55 (d, J = 8.4 Hz, 2H, Ar-H), 5.63 (d, J = 1.3 Hz, H-1), 4.15 (dd, J = 1.8, 3.1 Hz, H-2), 4.03 (dd, J = 3.4, 9.5 Hz, H-3), 3.94 (s, 2H, CH_2), 3.86-3.64 (m, 6H, H-4, H-5, H-6, CH_2); ^{13}C NMR (126 MHz, CD_3OD): δ = 153.21, 143.84, 136.41, 129.66, 129.18, 127.76, 127.70, 125.37, 118.64, 103.62 (Ar-C), 100.75 (C-1), 76.00 (C-5), 72.41 (C-3), 71.86 (C-2), 68.24 (C-4), 62.69 (C-6), 40.76 (CH_2); ESI-MS: m/z : Calcd for $\text{C}_{42}\text{H}_{37}\text{ClN}_3\text{O}_{12}\text{S}$ $[\text{M}+\text{H}]^+$: 842.2, found: 842.2.

3'-Chloro-*N*-(2-(2-(2-(3-(3',6'-dihydroxy-3-oxo-3*H*-spiro[isobenzofuran-1,9'-xanthen]-5-yl)thioureido)ethoxy)ethoxy)ethyl)-4'-(α -D-mannopyranosyloxy)-biphenyl-4-

carboxamide (24). Compound **21** (280 mg, 0.68 mmol) was dissolved in dry DMF (5 mL) under argon, then NHS (235 mg, 2.04 mmol) was added, followed by DIC (0.12 mL, 0.78 mmol) and the mixture was stirred at rt for 4 h, then Boc-PEG2-NH₂ (186 mg, 0.75 mmol) was added, and the mixture was stirred at rt under argon for 10 h. It was then slowly diluted with water and concentrated. The residue was purified by chromatography on silica gel (DCM/MeOH) to give *tert*-butyl (2-(2-(2-(3'-chloro-4'-(α -D-mannopyranosyloxy)-biphenyl-4-ylcarboxamido)ethoxy)ethoxy)ethyl)carbamate (300 mg, 0.468 mmol, 69%). Then, the carbamate was suspended in DCM (3 mL) and TFA (1 mL) was added dropwise at rt. After 30 min, the solvents were evaporated and the crude mixture was dissolved in $\text{CHCl}_3/\text{MeOH}$ (6:4, + 0.5% conc. NH_4OH) and transferred to a silica gel column, eluting with the same solvent mixture, to yield *N*-(2-(2-(2-aminoethoxy)ethoxy)ethyl)-3'-chloro-4'-(α -D-mannopyranosyl-

oxy)-biphenyl-4-carboxamide (228 mg, 90 %). A fraction of the amine (10 mg, 0.018 mmol) was dissolved in dry DMF (0.5 mL) and cooled to 0 °C. FITC (6.5 mg, 0.017 mmol) was added and the mixture was stirred for 1 h. The mixture was concentrated and the residue was purified by chromatography on silica (DCM/MeOH), to yield **24** (10 mg, 65%). ¹H NMR (500 MHz, CD₃OD): δ = 8.21 (d, J = 1.4 Hz, 1H, Ar-H), 7.88 (d, J = 8.3 Hz, 2H, Ar-H), 7.68 (d, J = 2.2 Hz, 2H, Ar-H), 7.63 (d, J = 8.3 Hz, 2H, Ar-H), 7.53 (dd, J = 2.2, 8.6 Hz, 1H, Ar-H), 7.43 (d, J = 8.7 Hz, 1H, Ar-H), 7.09 (d, J = 8.2 Hz, 1H, Ar-H), 6.68 (d, J = 2.3 Hz, 2H, Ar-H), 6.65 (dd, J = 2.6, 8.6 Hz, 2H, Ar-H), 6.53 (dd, J = 1.6, 8.7 Hz, 2H, Ar-H), 5.61 (d, J = 1.3 Hz, 1H, H-1), 4.14 (dd, J = 1.8, 3.2, Hz, 1H, H-2), 4.03 (dd, J = 3.4, 9.5 Hz, 1H, H-3), 3.93-3.53 (m, 16H), 3.37 (s, 2H, NCH₂), 1.30 (s, 2H, CH₂); ¹³C NMR (126 MHz, CD₃OD): δ = 170.01 (CO), 153.17, 143.72, 136.37, 134.37, 130.39, 129.69, 129.04, 127.78, 127.73, 125.35, 118.60, 103.60 (Ar-C), 100.72 (C-1), 75.97 (C-5), 72.41 (C-3), 71.86, 71.40, 70.59 (5C, C-2, OCH₂), 68.23 (C-4), 62.64 (C-6), 49.88, 45.49, 40.97 (CH₂); ESI-MS: m/z : Calcd for C₄₆H₄₅ClN₃O₁₄S [M+H]⁺: 930.2, found: 930.4.

Competitive Fluorescence Polarization Assay

Expression and purification of CRD of FimH. A recombinant protein consisting of the CRD of FimH linked to a 6His-tag via a thrombin cleavage site (FimH-CRD-Th-His₆) was expressed in *E. coli* strain HM125 and purified by affinity chromatography as previously described.⁴³

K_D determination of FITC-labeled ligands. The functionalized ligands (**23**, **24**) were prepared as a 10 mM stock solution in pure DMSO (Sigma Aldrich, Buchs, Switzerland). All further dilutions of compounds and FimH-CRD-Th-His₆ protein were prepared in assay buffer (20 mM HEPES, 150 mM NaCl, 50 µg/mL BSA, pH 7.4). BSA was added to the assay buffer

to prevent non-specific binding of protein to the plastic surface. Binding isotherms for the fluorescent ligands were obtained in direct binding studies by adding a constant concentration of ligand (final concentration 5 nM) and a linear dilution of protein (final concentration 0-100 nM) to a final volume of 200 μ L in 96-well, black, flat bottom NBSTM plates (Corning Inc., Corning, NY, USA). After incubating the plate for 24 h at rt with gentle shaking, the fluorescence polarization was measured with the SynergyTM H1 Hybrid Multi-Mode Microplate Reader (BioTek Instruments Inc., Winooski, VT, USA) with polarized excitation at 485 nm and emission measured at 528 nm through polarizing filters parallel and perpendicularly oriented to the incident polarized light. K_D values were determined by plotting the FP readout as a function of the protein concentration and applying the following single-site binding equation (Equation 1) that accounts for ligand depletion:

$$S_{obs} = S_F + (S_B - S_F) \cdot \left(\frac{C_P + C_L + K_D - \sqrt{(C_P + C_L + K_D)^2 - 4C_PC_L}}{2C_L} \right) \quad (1)$$

where S_{obs} is the observed signal from the ligand, S_F is the signal from free ligand, S_B is the signal from bound ligand, C_P is the total concentration of protein, and C_L is the total concentration of ligand.⁴⁹

K_D Determination of FimH Antagonists. The fluorescently labeled ligand **23** was used for the competitive fluorescence polarization assay. A linear dilution of non-labeled FimH antagonist with final concentrations ranging from 0-10 μ M was titrated into 96-well, black, flat bottom NBSTM plates (Corning Inc.) to a final volume of 200 μ L containing a constant concentration of protein (final concentration 25 nM) and FITC-labeled ligand which was fixed at a higher concentration in competitive binding assays than in direct binding experiments to obtain higher fluorescence intensities (final concentration 20 nM). Prior to measuring the fluorescence polarization, the plates were incubated on a shaker for 24 h at rt until the reaction reached equilibrium. The IC_{50} value was determined with Prism (GraphPad Software Inc., La

Jolla, CA, USA) by applying a standard four-parameter IC_{50} function. The obtained IC_{50} values were converted into their corresponding K_D values using the derivation of the Cheng-Prusoff equation.⁴⁵ This variation of the Cheng-Prusoff equation is applied to competition assays with tight-binding inhibitors, and includes terms to correct for ligand depletion effects. However, the K_D for antagonists having a higher affinity towards FimH than the labeled ligand could not be accurately determined.⁴⁵

Isothermal Titration Calorimetry (ITC)

All ITC experiments were performed with the FimH-CRD-Th-His₆ protein using a VP-ITC instrument from MicroCal, Inc. (Malvern Instruments, Worcestershire, UK) with a sample cell volume of 1.4523 mL. The measurements were performed with 0 to 5% DMSO at 25 °C, a stirring speed of 307 rpm, and 10 $\mu\text{cal s}^{-1}$ reference power. The protein samples were dialyzed in assay buffer prior to all experiments. Due to the high protein consumption of ITC, only the experiments for the reference compounds (**1**, **3** and **25**) were measured in duplicates. Compounds **1**, **3**, **9**, and **25** were measured in a direct fashion by titration of ligand (100-2,000 μM) into protein (8.6-55 μM) with injections of 3-8 μL at intervals of 10 min to ensure non-overlapping peaks. The quantity $c = Mt(0) K_D^{-1}$, where $Mt(0)$ is the initial macromolecule concentration, is of importance in titration microcalorimetry. The c -values of the direct titrations were below 1,000 and thus within the reliable range. For the compounds **10b-e**, **10g** and **10j** additional competitive ITC experiments were performed due to their high affinity resulting in c -values above 1,000 for direct titrations. These ligands (600 μM) were titrated into protein (30 μM), which was preincubated with compound **25** (300 μM) resulting in sigmoidal titration curves. Due to slow reaction kinetics, titration intervals of 20 min were used.

Baseline correction and peak integration was performed using the Origin 7 software (OriginLab, Northampton, MA, USA). An initial 2 μL injection was excluded from data analysis. Baseline subtraction and curve-fitting with the three variables N (concentration correction factor), K_D (dissociation constant), and ΔH° (change in enthalpy) was performed with the SEDPHAT software version 10.40 (National Institute of Health).⁸⁴ A global fitting analysis was performed for the competition titration (compounds **10b-e**, **10g** or **10j** competing for the protein binding site with compound **25**) and the direct titration of the competitor (compound **25** binding to protein) to fit for K_D . ΔH° and N were fitted from direct titrations of compounds **10b-e**, **10g** or **10j** into protein. For the compounds **3**, **9** and **25** binding to protein all variables could be determined from a global analysis of the direct titration.

The thermodynamic parameters were calculated with the following equation (Equation 2):

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = RT \ln K_D = -RT \ln K_A \quad (2)$$

where ΔG° , ΔH° , and ΔS° are the changes in free energy, enthalpy, and entropy of binding, respectively, T is the absolute temperature, and R is the universal gas constant (8.314 J mol⁻¹ K⁻¹). The 95 %-confidence intervals of the measurements were calculated for the two variables K_D and ΔH° with the 1-dimensional error surface projection within the SEDPHAT software.

Calculation of the Free Energy of Desolvation. The three dimensional representation for each of the aglycons (4-methoxy biphenyl scaffold, Figure 7) was built in the Maestro⁸⁵ modeling environment and the global minimum conformation was identified by performing 500 iterations of the mixed torsional/low-mode conformational sampling in combination with the OPLS-2005 force-field and the implicit solvent model (water) as implemented in the Macromodel 9.9.⁸⁶ The global minimum structures were used as input for the AMSOL 7.1 program⁸⁷ to obtain the free energy of desolvation ΔG_{des} (Table 5) with the SM5.4A solvation

model⁸⁸ and the AM1⁸⁹ level of theory (used keywords “AM1 SM5.4A SOLVNT=WATER TRUES”).

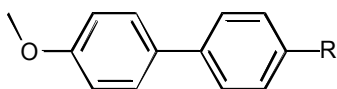


Figure 7. The 4-methoxy biphenyl scaffold of aglycons.

Table 5. Aqueous free energy of desolvation.

R	ΔG_{des} [kJ/mol]
<i>neutral</i>	
H	15.6
CONHCH ₃	39.9
COOCH ₃	23.0
SO ₂ NHCH ₃	65.5
SO ₂ CH ₃	56.4
4-morpholine amide	45.3
CN	22.0
<i>deprotonated</i>	
COO ⁻	298.2
SO ₂ -N ⁻ -Me	342.0

Determination of the MAC₉₀ by flow cytometry

The MAC₉₀ was determined in principle as in the previously published flow cytometry assay,⁷⁹ but with some modifications. The human epithelial bladder carcinoma cell line 5637 (DSMZ, Braunschweig, Germany) was grown in RPMI 1640 medium, supplemented with 10% fetal

calf serum (FCS), 100 U/mL penicillin and 100 µg/mL streptomycin at 37 °C, 5% CO₂. All solutions were purchased from Invitrogen (Basel, Switzerland). The cells were subcultured 1:6 twice per week [using Trypsin/EDTA (Sigma-Aldrich) for the detachment]. Two days before infection, 1.8×10^5 cells were seeded in each well of a 24-well plate in RPMI 1640 containing 10% FCS without antibiotics. The cell density was approximately $3\text{--}5 \times 10^5$ cells/well at the assay day.

For infection, the GFP-expressing clinical *E. coli* isolate UTI89⁹⁰ (UTI89 wt) and the GFP-expressing FimA-H knock-out strain UTI89 Δ *fimA-H* were used (strains were provided by Prof. Urs Jenal, Biocenter, University of Basel, Switzerland).⁷⁹ Bacteria were cultivated at 37 °C in 10 mL Luria-Bertani (LB) broth (Becton, Dickinson and Company) overnight, harvested by centrifugation (3,800 rpm, 10 min) and washed three times in phosphate buffered saline (PBS, Sigma-Aldrich) and a bacterial solution of OD₆₀₀ of 0.75 in RPMI + 10% FCS was prepared. For the determination of the MAC₉₀ value, the IC₉₀, linear dilutions of the FimH-antagonist were prepared in 5% DMSO and PBS. Bacteria and antagonists were pre-incubated for 10 min at 37 °C, before cells were infected with either only 200 µL bacterial solution of UTI89 or UTI89 Δ *fimA-H* (positive and negative controls), or 225 µL of the pre-incubated bacteria-antagonist mixture. Infection lasted for 1.5 h, during this time infected cells were incubated at 37 °C. Then, cells were washed with PBS and detached from wells by the addition of 150 µL trypsin and incubation at 37 °C for 10 min, before flushing from wells PBS containing 2% FCS and transferred to tubes. To dilute the trypsin, cells were centrifuged at 13000 rpm, 1 min, 600 µL of the supernatant was discarded and the pellet was re-suspended in the remaining 300 µL PBS containing 2% FCS. Samples were stored on ice until measurement. Before analysis with the flow cytometer (Becton Dickinson, FACSCanto II), the samples were gently mixed and filtered using a 35 µm nylon mesh (Corning Life Sciences) to

prevent cellular aggregation. Cells were gated with linear scaling for side scatter (SSC) and forward scatter (FSC) and GFP intensity of live cells was evaluated. IC₉₀ values were determined by plotting the concentration of the antagonist in a logarithmic mode versus the mean fluorescence intensity (MFI) of living cells and by fitting a dose response curve (variable slope, four parameters) with the prism software (GraphPad Prism).

X-ray analysis of the antagonists 10e and 10j co-crystallized with FimH-CRD

FimH-CRD-10e co-crystallization. Initial FimH-CRD (18 mg/mL in 20 mM HEPES pH 7.4) crystals were obtained in complex with 4-(5-nitroindolin-1-yl)phenyl α -D-mannopyranoside (5 mM).²³ Crystals were grown in sitting-drop vapor diffusion at 20 °C with 200 nL of protein-antagonist mixture together with 200 nL precipitant solution in well D3 (0.2 M sodium phosphate monobasic monohydrate, 20% w/v PEG 3,350) of the PEG/Ion HTTM screen (Hampton Research, CA, USA). Cubic crystals appeared within one week, which served as cross-seeding crystals. A solution of FimH-CRD (20 mg/mL) and **10e** (5 mM) was mixed with 0.2 M sodium phosphate monobasic monohydrate, 20% w/v PEG 400 with 0.5 μ L of each solution. Streak-seeding was performed after one day of incubation. Cubic FimH-CRD-**10e** crystals formed within 24 h. Crystals were flash cooled to 100 K with perfluoropolyether cryo oil (Hampton Research, CA, USA) as cryoprotectant. Data was collected with synchrotron radiation ($\lambda = 0.99999$ Å) at the PXIII beamline, Swiss Light Source, Switzerland.

FimH-CRD-10j co-crystallization. Co-crystals were initially grown in sitting-drop vapor diffusion at 20 °C with 0.5 μ L of a mixture of FimH-CRD (20 mg/mL) and **10j** (5 mM) together with 0.5 μ L of 0.1 M HEPES pH 7.5, 2 M ammonium sulfate. Plate-like crystals formed within two weeks and were used as seeds for subsequent crystallization. Diffraction

quality crystals were grown by streak-seeding in 0.5 μL of FimH-CRD (10 mg/mL) with **10j** (2.5 mM) and 0.5 μL of 0.1 M HEPES pH 7.5, 1.25 M ammonium sulfate. The drops were covered with perfluoropolyether cryo oil prior to flash cooling to 100 K. Data was collected with synchrotron radiation ($\lambda = 1.00003 \text{ \AA}$) at the PXIII beamline, Swiss Light Source, Switzerland.

Structure Determination and Refinement. Data were indexed and integrated with the XDS package⁹¹ for the FimH-CRD-**10e** co-crystal structure, and with mosflm⁹² for the FimH-CRD-**10j** co-crystal structure. Scaling was performed with XDS and SCALA included in the CCP4 suite, respectively.⁹³ Structures were solved by molecular replacement with PHASER⁹⁴ using the FimH-CRD-butyl α -D-mannopyranoside complex (PDB code 1UWF) as search model. The structures were iteratively built using the COOT software⁹⁵ and refined with the PHENIX software.⁹⁶ Geometric restraints for **10e** and **10j** were generated with PRODRG.⁹⁷ The models were validated using molprobity.⁹⁸ Residues 113-115 were not modeled in the **10e** structure due to disorder. Furthermore, the ligand was modeled in two possible conformations. For both ligands, electron density is reduced on the second aromatic ring due to flexibility of the ligand.

Table 6. Data Collection and Refinement Statistics for FimH-CRD-**10e** and FimH-CRD-**10j** co-crystals.

	FimH-CRD- 10e	FimH-CRD- 10j
PDB code	4CSS	4CST
Space group	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁
No. of molecules in the asymmetric unit	1	1
Cell dimensions		
<i>a</i> , <i>b</i> , <i>c</i> (Å)	48.38, 56.23, 61.59	48.84, 55.89, 61.00
<i>a</i> , <i>b</i> , <i>c</i> (°)	90, 90, 90	90, 90, 90

Data Collection		
Beamline	Swiss Light Source PXIII	Swiss Light Source PXIII
Resolution range (Å) ^a	30.0 - 1.07 (1.13 – 1.07)	23.5 – 1.10 (1.12 – 1.10)
Unique observations ^a	72000 (9354)	66470 (2500)
Average multiplicity ^a	10.9 (3.7)	5.4 (2.4)
Completeness (%)	96.1 (78.0)	97.2 (76.5)
R_{merge} ^a	0.056 (0.57)	0.051 (0.305)
Mean I / $\sigma(I)$ ^a	21.5 (2.22)	15.5 (2.9)
Refinement		
Resolution range (Å)	15.7 – 1.07	23.5 – 1.10
R , R_{free}	11.2, 13.2	11.4, 13.0
Rms deviation from ideal bond length (Å)	0.010	0.010
Rms deviation from ideal bond angle (deg)	1.170	1.420

^aValues in parentheses are for highest-resolution shell.

Physicochemical and *in vitro* pharmacokinetic studies.

Materials. Dimethyl sulfoxide (DMSO), 1-propanol, 1-octanol, Dulbecco's Modified Eagle's Medium (DMEM) - high glucose, L-glutamine solution, penicillin-streptomycin solution, Dulbecco's Phosphate Buffered Saline (DPBS), trypsin-EDTA solution, magnesium chloride hexahydrate, and reduced nicotinamide adenine dinucleotide phosphate (NADPH) were purchased from Sigma-Aldrich. MEM nonessential amino acid (MEM-NEAA) solution, fetal bovine serum (FBS), and DMEM without sodium pyruvate and phenol red were bought from Invitrogen (Carlsbad, CA, USA). PRISMA HT universal buffer, GIT-0 Lipid Solution, and

Acceptor Sink Buffer were ordered from pIon (Woburn, MA, USA). Human plasma was bought from Biopredic (Rennes, France) and acetonitrile (MeCN) and methanol (MeOH) from Acros Organics (Geel, Belgium). Pooled male rat liver microsomes were purchased from BD Bioscience (Franklin Lakes, NJ, USA). Tris(hydroxymethyl)-aminomethane (TRIS) was obtained from AppliChem (Darmstadt, Germany). The Caco-2 cells were kindly provided by Prof. G. Imanidis, FHNW, Muttentz, and originated from the American Type Culture Collection (Rockville, MD, USA).

pK_a. The pK_a values were determined as described elsewhere.⁹⁹ In brief, the pH of a sample solution was gradually changed and the chemical shift of protons adjacent to ionizable centers was monitored by ¹H nuclear magnetic resonance (NMR) spectroscopy. The shift was plotted against the pH of the respective sample, and the pK_a was read out from the inflection point of the resulting sigmoidal curve.

Log D_{7.4}. The in silico prediction tool ALOGPS¹⁰⁰ was used to estimate log *P* values of the compounds. Depending on these values, the compounds were classified into three categories: hydrophilic compounds (log *P* below zero), moderately lipophilic compounds (log *P* between zero and one) and lipophilic compounds (log *P* above one). For each category, two different ratios (volume of 1-octanol to volume of buffer) were defined as experimental parameters (Table 7).

Table 7. Compound classification based on estimated log *P* values.

Compound type	log <i>P</i>	ratio (1-octanol / buffer)
hydrophilic	< 0	30:140, 40:130
moderately lipophilic	0 - 1	70:110, 110:70
lipophilic	> 1	3:180, 4:180

Equal amounts of phosphate buffer (0.1 M, pH 7.4) and 1-octanol were mixed and shaken vigorously for 5 min to saturate the phases. The mixture was left until separation of the two phases occurred, and the buffer was retrieved. Stock solutions of the test compounds were diluted with buffer to a concentration of 1 μ M. For each compound, six determinations, that is, three determinations per 1-octanol/buffer ratio, were performed in different wells of a 96-well plate. The respective volumes of buffer containing analyte (1 μ M) were pipetted to the wells and covered by saturated 1-octanol according to the chosen volume ratio. The plate was sealed with aluminum foil, shaken (1,350 rpm, 25 °C, 2 h) on a Heidolph Titramax 1000 plate-shaker (Heidolph Instruments GmbH & Co. KG, Schwabach, Germany) and centrifuged (2,000 rpm, 25 °C, 5 min, 5804 R Eppendorf centrifuge, Hamburg, Germany). The aqueous phase was transferred to a 96-well plate for analysis by LC-MS.

The $\log D_{7.4}$ coefficient was calculated from the 1-octanol/buffer ratio (o:b), the initial concentration of the analyte in buffer (1 μ M), and the concentration of the analyte in buffer (c_B) with Equation 3:

$$\log D_{7.4} = \log \left(\frac{1\mu M - c_B}{c_B} \times \frac{1}{o:b} \right) \quad (3)$$

Aqueous Solubility. Solubility was determined in a 96-well format using the μ SOL Explorer solubility analyzer (pIon, version 3.4.0.5). For each compound, measurements were performed at pH 3.0 and 7.4 in triplicates. For this purpose, six wells of a deep well plate, that is, three wells per pH value, were filled with 300 μ L of PRISMA HT universal buffer, adjusted to pH 3.0 or 7.4 by adding the requested amount of NaOH (0.5 M). Aliquots (3 μ L) of a compound stock solution (10-50 mM in DMSO) were added and thoroughly mixed. The final sample concentration was 0.1-0.5 mM, the residual DMSO concentration was 1.0% (v/v) in the buffer

solutions. After 15 h, the solutions were filtrated (0.2 μm 96-well filter plates) using a vacuum to collect manifold (Whatman Ltd., Maidstone, UK) to remove the precipitates. Equal amounts of filtrate and 1-propanol were mixed and transferred to a 96-well plate for UV/Vis detection (190 to 500 nm, SpectraMax 190). The amount of material dissolved was calculated by comparison with UV/Vis spectra obtained from reference samples, which were prepared by dissolving compound stock solution in a 1:1 mixture of buffer and 1-propanol (final concentrations 0.017-0.083 mM).

Parallel Artificial Membrane Permeation Assay (PAMPA). Effective permeability ($\log P_e$) was determined in a 96-well format with the PAMPA.⁶⁰ For each compound, measurements were performed at pH 5.0 and 7.4 in quadruplicates. Eight wells of a deep well plate, that is, four wells per pH-value, were filled with 650 μL of PRISMA HT universal buffer adjusted to pH 5.0 or 7.4 by adding the requested amount of NaOH (0.5 M). Samples (150 μL) were withdrawn from each well to determine the blank spectra by UV/Vis-spectroscopy (190 to 500 nm, SpectraMax 190). Then, analyte dissolved in DMSO was added to the remaining buffer to yield 50 μM solutions. To exclude precipitation, the optical density was measured at 650 nm, with 0.01 being the threshold value. Solutions exceeding this threshold were filtrated. Afterwards, samples (150 μL) were withdrawn to determine the reference spectra. Further 200 μL were transferred to each well of the donor plate of the PAMPA sandwich (pIon, P/N 110163). The filter membranes at the bottom of the acceptor plate were infused with 5 μL of GIT-0 Lipid Solution, and 200 μL of Acceptor Sink Buffer was filled into each acceptor well. The sandwich was assembled, placed in the GutBoxTM, and left undisturbed for 16 h. Then, it was disassembled and samples (150 μL) were transferred from each donor and acceptor well to UV-plates for determination of the UV/Vis spectra. Effective permeability ($\log P_e$) was calculated from the compound flux deduced from the spectra, the filter area, and the initial

sample concentration in the donor well with the aid of the PAMPA Explorer Software (pIon, version 3.5).

Colorectal Adenocarcinoma (Caco-2) Cell Permeation Assay. Caco-2 cells were cultivated in tissue culture flasks (BD Biosciences) with DMEM high glucose medium, containing L-glutamine (2 mM), nonessential amino acids (0.1 mM), penicillin (100 U/mL), streptomycin (100 µg/mL), and fetal bovine serum (10%). The cells were kept at 37 °C in humidified air containing 5% CO₂, and the medium was changed every second day. When approximately 90% confluence was reached, the cells were split in a 1:10 ratio and distributed to new tissue culture flasks. At passage numbers between 60 and 65, they were seeded at a density of 5.3×10^5 cells per well to Transwell 6-well plates (Corning Inc.) with 2.5 mL of culture medium in the basolateral and 1.8 mL in the apical compartment. The medium was renewed on alternate days. Permeation experiments were performed between days 19 and 21 post seeding. Prior to the experiment, the integrity of the Caco-2 monolayers was evaluated by measuring the transepithelial electrical resistance (TEER) with an Endohm tissue resistance instrument (World Precision Instruments Inc., Sarasota, FL, USA). Only wells with TEER values higher than $250 \Omega \text{ cm}^2$ were used. Experiments were performed in the apical-to-basolateral (absorptive) and basolateral-to-apical (secretory) directions in triplicates. Transport medium (DMEM without sodium pyruvate and phenol red) was withdrawn from the donor compartments of three wells and replaced by the same volume of compound stock solution (10 mM in DMSO) to reach an initial sample concentration of 62.5 µM. The Transwell plate was then shaken (600 rpm, 37 °C) on a Heidolph Titramax 1000 plate-shaker. Samples (40 µL) were withdrawn from the donor and acceptor compartments 30 min after initiation of the experiment and the compound concentrations were determined by LC-MS (see below). Apparent permeability (P_{app}) was calculated according to Equation 4:

$$P_{\text{app}} = \frac{dQ}{dt} \times \frac{1}{A \times c_0} \quad (4)$$

where dQ/dt is the compound flux (mol s^{-1}), A is the surface area of the monolayer (cm^2), and c_0 is the initial concentration in the donor compartment (mol cm^{-3}).⁶⁰ After the experiment, TEER values were assessed again for each well and results from wells with values below $250 \Omega \text{ cm}^2$ were discarded.

Plasma Protein Binding (PPB). PPB was determined in a 96-well format using a high throughput dialysis block (HTD96b; HTDialysis LCC, Gales Ferry, CT, USA). For each compound, measurements were performed in triplicate. Dialysis membranes (MWCO 12-14 K; HTDialysis LCC) were hydrated according to the instructions of the manufacturer and placed into the dialysis block. Human plasma was centrifuged (5,800 rpm, 5 °C, 10 min), the pH of the supernatant (without floating plasma lipids) was adjusted to 7.4 by adding the requested amount of HCl (4 M), and analyte was added to yield a final concentration of 10 μM . Equal volumes (150 μL) of plasma containing the analyte or TRIS-HCl buffer (0.1 M, pH 7.4) were transferred to the compartments separated by the dialysis membrane. The block was covered with a sealing film and left undisturbed (5 h, 37 °C). Afterwards, samples (90 μL) were withdrawn from the buffer compartments and diluted with plasma (10 μL). From the plasma compartments, samples (10 μL) were withdrawn and diluted with TRIS-HCl buffer (90 μL). The solutions were further diluted with ice-cooled MeCN (300 μL) to precipitate the proteins and centrifuged (3600 rpm, 4 °C, 10 min). The supernatants (50 μL) were retrieved, and the analyte concentrations were determined by LC-MS (see below). The fraction bound (f_b) was calculated as follows (Equation 5):

$$f_b = 1 - \frac{c_b}{c_p} \quad (5)$$

where c_b is the concentration of the analyte withdrawn from the buffer compartment before dilution and c_p is the concentration in the plasma compartment. The values were accepted if the recovery of analyte was between 80 and 120% of the initial amount.

Cytochrome P450-mediated metabolism. Incubations consisted of pooled male rat liver microsomes (0.5 mg microsomal protein/mL), test compound (2 μ M), $MgCl_2$ (2 mM), and NADPH (1 mM) in a total volume of 300 μ L TRIS-HCl buffer (0.1 M, pH 7.4) and were performed in a 96-well plate on a Thermomixer Comfort (Eppendorf). Compounds and microsomes were preincubated (37 °C, 700 rpm, 10 min) before NADPH was added. Samples (50 μ L) at $t = 0$ min and after an incubation time of 5, 10, 20, and 30 min were quenched with 150 μ L of ice-cooled MeOH, centrifuged (3,600 rpm, 4 °C, 10 min), and 80 μ L of supernatant was transferred to a 96-well plate for LC-MS analysis (see below). The metabolic half-life ($t_{1/2}$) was calculated from the slope of the linear regression from the log percentage remaining compound versus incubation time relationship. Control experiments without NADPH were performed in parallel.

LC-MS measurements. Analyses were performed using an 1100/1200 Series HPLC System coupled to a 6410 Triple Quadrupole mass detector (Agilent Technologies, Inc., Santa Clara, CA, USA) equipped with electrospray ionization. The system was controlled with the Agilent MassHunter Workstation Data Acquisition software (version B.01.04). The column used was an Atlantis® T3 C18 column (2.1 \times 50 mm) with a 3 μ m particle size (Waters Corp., Milford, MA, USA). The mobile phase consisted of eluent A: H_2O containing 0.1% formic acid (for **10a-f**, **h-i**), or 10 mM ammonium acetate, pH 5.0 in 95:5, H_2O /MeCN (for **10g**, **j**); and eluent B: MeCN containing 0.1% formic acid. The flow rate was maintained at 0.6 mL/min. The gradient was ramped from 95% A/5% B to 5% A/95% B over 1 min, and then hold at 5%

A/95% B for 0.1 min. The system was then brought back to 95% A/5% B, resulting in a total duration of 4 min. MS parameters such as fragmentor voltage, collision energy, polarity were optimized individually for each analyte, and the molecular ion was followed for each compound in the multiple reaction monitoring mode. The concentrations of the analytes were quantified by the Agilent Mass Hunter Quantitative Analysis software (version B.01.04).

***In Vivo* Pharmacokinetic Studies.**

Materials. DMSO and PBS were purchased from Sigma-Aldrich. The 96-well plates were bought from Agilent Technologies (0.5 mL, polypropylene). The gavage was obtained from Fine Science (Heidelberg, Germany) and the syringes (BD Micro Fine, U-100 Insuline, 30 G) and needles (BD Microlance 3, 25 G) from Becton Dickinson (USA, Ireland) and Henke Sass Wolf in Germany (Soft-Ject, 1 mL syringes).

Animals. Female C3H/HeN mice weighing between 19 and 25 g were obtained from Charles River Laboratories (Sulzfeld, Germany) and were housed three or four per cage. The mice were kept under specific pathogen-free conditions in the Animal House of the Department of Biomedicine, University Hospital of Basel, and animal experimentation guidelines according to the regulations of the Swiss veterinary law were followed. After 7 d of acclimatization, 9-10 week old mice were used for the studies. Animals had free access to chow and water at any time and were kept in a 12h/12h light/dark cycle. For administration volumes and sampling the good practice guidelines were followed.¹⁰¹

Pharmacokinetic studies. The single-dose studies for the first experiment set were performed by intravenous application of FimH antagonists at a dosage of 50 mg/kg body weight, followed by plasma and urine sampling. Antagonists were diluted in PBS for injection into the tail vein.

Blood and urine samples (10 μ L) were taken at 6 and 30 min, and 1, 2, 4, 6, and 8 h after injection. For the PK studies with **10j**, the antagonist was dissolved in PBS with 5 % DMSO and injected into the tail vein (0.625 mg/kg) or given orally (1.25 mg/kg) using a gavage. Blood and urine were sampled (10 μ L) after 7, 13, 20, 30, 45 min, and 1, 1.5, 2, 2.5, 3, 4, 6, 8, and 24 h. Both, blood and urine samples, were directly diluted after sampling with MeOH to precipitate the proteins and centrifuged for 11 min at 13,000 rpm. The supernatants were transferred to a 96-well plate and the analyte concentrations were determined by LC-MS (see above).

Infection study. For all infection studies, the drinking water of the mice was replaced by water containing 5% glucose (monohydrate from AppliChem, BioChemica), three days before the start of the experiment. **10j** was dosed at 1.25 mg/kg (in 5% DMSO and PBS) and 10 mg/kg (5% DMSO in PBS containing 1% Tween 80) applied orally to 6 and 4 mice, respectively, as described in the pharmacokinetic studies, 40 min prior to infection. Ciprofloxacin was dosed with 8 mg/kg, which would correspond to a human dose of 400 mg,⁸¹ subcutaneously 10 min prior to infection with UTI89 to 4 mice. Control values resulted from the infection of 11 mice. Before infection, remaining urine in the bladder was expelled by gentle pressure on the abdomen. Mice were anaesthetised in 2.5 vol% isoflurane/oxygen mixture (Attane, Minrad Inc, Buffalo, NY, USA) and placed on their back. Infection was performed transurethrally using a polyethylene catheter (Intramedic polyethylene tubing, inner diameter 0.28 mm, outer diameter 0.61 mm, Beckton Dickinson, Allschwil, Switzerland), on a syringe (Hamilton Gastight Syringe 50 μ L, removable 30G needle, BGB Analytik AG, Boeckten, Switzerland). After gentle insertion of the catheter into the bladder, 50 μ L of bacterial suspension of UTI89 (5.5×10^9 to 2.25×10^{10} CFU/mL) was slowly injected. This corresponded to approximately 10^7 - 10^8 CFU per mouse. Mice were killed by CO₂ three hours after inoculation and bladder

and kidneys were aseptically removed. Organs were homogenized in 1 mL PBS using a tissue lyser (Retsch, Haan, Germany). Serial dilutions of bladder and kidneys were plated on Levine Eosin Methylene Blue Agar plates (Beckton Dickinson, France) and CFU were counted after overnight incubation at 37 °C.

Authors Information

Corresponding author

*Tel: +41 61 267 15 51. Fax: +41 61 267 15 52. E-mail: beat.ernst@unibas.ch

Author contributions

These authors contributed equally to the project

Notes

The authors declare no competing financial interest.

Acknowledgement

The authors thank Prof. Dr. med. Radek Skoda, Department of Biomedicine, University Hospital Basel, Switzerland, for giving us access to the animal facility. The financial support by the Swiss National Science Foundation (SNF interdisciplinary grant K-32K1-120904) is gratefully acknowledged.

References

- (1) Foxman, B.; Barlow, R.; D’Arcy, H.; Gillespie, B.; Sobel, J. D. Urinary tract infection: self-reported incidence and associated costs. *Ann. Epidemiol.* **2000**, *10*, 509-515.
- (2) Ronald, A. The etiology of urinary tract infection: traditional and emerging pathogens. *Am. J. Med.* **2002**, *113 Suppl. 1A*, 14S-19S.

- (3) Fihn, S. D. Acute uncomplicated urinary tract infection in women. *N. Engl. J. Med.* **2003**, *349*, 259-266.
- (4) Hooton, T. M.; Besser, R.; Foxman, B.; Fritsche, T. R.; Nicolle, L. E. Acute uncomplicated cystitis in an era of increasing antibiotic resistance: a proposed approach to empirical therapy. *Clin. Infect. Dis.* **2004**, *39*, 75-80.
- (5) Sanchez, G. V.; Master, R. N.; Karlowsky, J. A.; Bordon, J. M. In vitro antimicrobial resistance of urinary *Escherichia coli* isolates among U. S. outpatients from 2000 to 2010. *Antimicrob. Agents Chemother.* **2012**, *56*, 2181-2183.
- (6) Clatworthy, A. E.; Pierson, E.; Hung, D. T. Targeting virulence: a new paradigm for antimicrobial therapy. *Nature Chem. Biol.* **2007**, *3*, 541-548.
- (7) Mulvey, M. A.; Schilling, J. D.; Martinez, J. J.; Hultgren, S. J. Bad bugs and beleaguered bladders: interplay between uropathogenic *Escherichia coli* and innate host defenses. *Proc. Natl. Acad. Sci. U S A* **2000**, *97*, 8829-8835.
- (8) Schilling, J. D.; Mulvey, M. A.; Hultgren, S. J. Structure and function of *Escherichia coli* type 1 pili: new insight into the pathogenesis of urinary tract infections. *J. Infect. Dis.* **2001**, *183 Suppl. 1*, S36-40.
- (9) Wiles, T. J.; Kulesus, R. R.; Mulvey, M. A. Origins and virulence mechanisms of uropathogenic *Escherichia coli*. *Exp. Mol. Pathol.* **2008**, *85*, 11-19.
- (10) Capitani, G.; Eidam, O.; Glockshuber, R.; Grütter, M. G. Structural and functional insights into the assembly of type 1 pili from *Escherichia coli*. *Microbes Infect.* **2006**, *8*, 2284-2290.
- (11) Le Trong, I.; Aprikian, P.; Kidd, B. A.; Forero-Shelton, M.; Tchesnokova, V.; Rajagopal, P.; Rodriguez, V.; Interlandi, G.; Klevit, R.; Vogel, V.; Stenkamp, R. E.;

- Sokurenko, E. V.; Thomas, W. E. Structural basis for mechanical force regulation of the adhesin FimH via finger trap-like β sheet twisting. *Cell* **2010**, *141*, 645-655.
- (12) Sharon, N. Carbohydrates as future anti-adhesion drugs for infectious diseases. *Biochim. Biophys. Acta*. **2006**, *1760*, 527-537.
- (13) Firon, N.; Itzhak, O.; Sharon, N. Interaction of mannose-containing oligosaccharides with the fimbrial lectin of *Escherichia coli*. *Biochem. Biophys. Res. Commun.* **1982**, *105*, 1426-1432.
- (14) Firon, N.; Ofek, I.; Sharon, N. Carbohydrate specificity of the surface lectins of *Escherichia coli*, *Klebsiella pneumoniae*, and *Salmonella typhimurium*. *Carbohydr. Res.* **1983**, *120*, 235-249.
- (15) Bouckaert, J.; Berglund, J.; Schembri, M.; De Genst, E.; Cools, L.; Wuhler, M.; Hung, C.-S.; Pinkner, J.; Slättegård, R.; Zavialov, A.; Choudhury, D.; Langermann, S.; Hultgren, S. J.; Wyns, L.; Klemm, P.; Oscarson, S.; Knight, S. D.; De Greve, H. Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.* **2005**, *55*, 441-455.
- (16) Firon, N.; Ashkenazi, S.; Mirelman, D.; Ofek, I.; Sharon, N. Aromatic alpha-glycosides of mannose are powerful inhibitors of the adherence of type 1 fimbriated *Escherichia coli* to yeast and intestinal epithelial cells. *Infect. Immun.* **1987**, *55*, 472-476.
- (17) Sperling, O.; Fuchs, A.; Lindhorst, T. K. Evaluation of the carbohydrate recognition domain of the bacterial adhesin FimH. Design, synthesis and binding properties of mannoside ligands. *Org. Biomol. Chem.* **2006**, *4*, 3913-3922.
- (18) Han, Z.; Pinkner, J. S.; Ford, B.; Obermann, R.; Nolan, W.; Wildman, S. A.; Hobbs, D.; Ellenberger, T.; Cusumano, C. K.; Hultgren, S. J.; Janetka, J. W. Structure-based

- drug design and optimization of mannoside bacterial FimH antagonists. *J. Med. Chem.* **2010**, *53*, 4779-4792.
- (19) Klein, T.; Abgottspon, D.; Wittwer, M.; Rabbani, S.; Herold, J.; Jiang, X.; Kleeb, S.; Lüthi, C.; Scharenberg, M.; Bezençon, J.; Gubler, E.; Pang, L.; Smiesko, M.; Cutting, B.; Schwardt, O.; Ernst, B. FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J. Med. Chem.* **2010**, *53*, 8627-8641.
- (20) Cusumano, C. K.; Pinkner, J. S.; Han, Z.; Greene, S. E.; Ford, B. A.; Crowley, J. R.; Henderson, J. P.; Janetka, J. W.; Hultgren, S. J. Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci. Transl. Med.* **2011**, *3*, 109ra115.
- (21) Han, Z.; Pinkner, J. S.; Ford, B.; Chorell, E.; Crowley, J. M.; Cusumano, C. K.; Campbell, S.; Henderson, J. P.; Hultgren, S. J.; Janetka, J. W. Lead optimization studies on FimH antagonists: discovery of potent and orally bioavailable *ortho*-substituted biphenyl mannosides. *J. Med. Chem.* **2012**, *55*, 3945-3959.
- (22) Pang, L.; Kleeb, S.; Lemme, K.; Rabbani, S.; Scharenberg, M.; Zalewski, A.; Schädler, F.; Schwardt, O.; Ernst, B. FimH antagonists: structure-activity and structure-property relationships for biphenyl α -D-mannopyranosides. *ChemMedChem.* **2012**, *7*, 1404-1422.
- (23) Jiang, X.; Abgottspon, D.; Kleeb, S.; Rabbani, S.; Scharenberg, M.; Wittwer, M.; Haug, M.; Schwardt, O.; Ernst, B. Anti-adhesion therapy for urinary tract infections – a balanced PK/PD profile proved to be key for success. *J. Med. Chem.* **2012**, *55*, 4700-4713.
- (24) Schwardt, O.; Rabbani, S.; Hartmann, M.; Abgottspon, D.; Wittwer, M.; Kleeb, S.; Zalewski, A.; Smiesko, M.; Cutting, B.; Ernst, B. Design, synthesis and biological

- evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.* **2011**, *19*, 6454-6473.
- (25) Brument, S.; Sivignon, A.; Dumych, T. I.; Moreau, N.; Roos, G.; Guérardel, Y.; Chalopin, T.; Deniaud, D.; Bilyy, R. O.; Darfeuille-Michaud, A.; Bouckaert, J.; Gouin, S. G. Thiazolylaminomannosides as potent antiadhesives of type 1 piliated *Escherichia coli* isolated from Crohn's disease patients. *J. Med. Chem.* **2013**, *56*, 5395-5406.
- (26) Lindhorst, T. K.; Kieburg, C.; Krallmann-Wenzel, U. Inhibition of the type 1 fimbriae-mediated adhesion of *Escherichia coli* to erythrocytes by multiantennary D-mannosyl clusters: The effect of multivalency. *Glycoconj. J.* **1998**, *15*, 605-613.
- (27) Nagahori, N.; Lee, R. T.; Nishimura, S.-L.; Pagé, S.; Roy, R.; Lee, Y. C. Inhibition of adhesion of type 1 fimbriated *Escherichia coli* to highly mannosylated ligands. *ChemBioChem* **2002**, *3*, 836-844.
- (28) Appeldoorn, C. C. M.; Joosten, J. A. F.; Maate, F. A.; Dobrindt, U.; Hacker, J.; Liskamp, R. M. J.; Khan, A. S.; Pieters, R. J. Novel multivalent mannose compounds and their inhibition of the adhesion of type 1 fimbriated uropathogenic *E. coli*. *Tetrahedron Asym.* **2005**, *16*, 361-372.
- (29) Patel, A.; Lindhorst, T. K. A modular approach for the synthesis of oligosaccharide mimetics. *Carbohydr. Res.* **2006**, *341*, 1657-1668.
- (30) Touaibia, M.; Wellens, A.; Shiao, T. C.; Wang, Q.; Sirois, S.; Bouckaert, J.; Roy, R. Mannosylated G(0) dendrimers with nanomolar affinities to *Escherichia coli* FimH. *ChemMedChem* **2007**, *2*, 1190-1201.
- (31) Durka, M.; Buffet, K.; Iehl, J.; Holler, M.; Nierengarten, J.-F.; Taganna, J.; Bouckaert, J.; Vincent, S. P. The functional valency of dodecamannosylated fullerenes with

- Escherichia coli* FimH - towards novel bacterial antiadhesives. *Chem. Commun.* **2011**, 47, 1321-1323.
- (32) Bouckaert, J.; Li, Z.; Xavier, C.; Almant, M.; Caveliers, V.; Lahoutte, T.; Weeks, S. D.; Kovensky, J.; Gouin, S. G. Heptyl α -D-mannosides grafted on a β -cyclodextrin core to interfere with *Escherichia coli* adhesion: An *in vivo* multivalent effect. *Chem. Eur. J.* **2013**, 19, 7847-7855.
- (33) Scharenberg, M.; Schwardt, O.; Rabbani, S.; Ernst, B. Target selectivity of FimH antagonists. *J. Med. Chem.* **2012**, 55, 9810-9816.
- (34) Choudhury, D.; Thompson, A.; Stojanoff, V.; Langermann, S.; Pinkner, J.; Hultgren, S. J.; Knight, S. D. X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* **1999**, 285, 1061-1066.
- (35) Hung, C.-S.; Bouckaert, J.; Hung, D.; Pinkner, J.; Widberg, C.; DeFusco, A.; Auguste, C. G.; Strouse, R.; Langermann, S.; Waksman, G.; Hultgren, S. J. Structural basis of tropism of *Escherichia coli* to the bladder drug in urinary tract infection. *Mol. Microbiol.* **2002**, 44, 903-915.
- (36) Wellens, A.; Garofalo, C.; Nguyen, H.; Van Gerven, N.; Slättegård, R.; Henalsteens, J.-P.; Wyns, L.; Oscarson, S.; De Greve, H.; Hultgren, S. J.; Bouckaert, J. Intervening with urinary tract infections using anti-adhesives based on the crystal structure of the FimH-oligomannose-3 complex. *PLoS One* **2008**, 3, e2040.
- (37) Wellens, A.; Lahmann, M.; Touaibia, M.; Vaucher, J.; Oscarson, S.; Roy, R.; Remaut, H.; Bouckaert, J. The tyrosine gate as a potential entropic lever in the receptor-binding site of the bacterial adhesin FimH. *Biochemistry* **2012**, 51, 4790-4799.
- (38) Totsika, M.; Kostakioti, M.; Hannan, T. J.; Upton, M.; Beatson, S. A.; Janetka, J. W.; Hultgren, S. J.; Schembri, M. A. A FimH inhibitor prevents acute bladder infection and

- treats chronic cystitis caused by multidrug-resistant uropathogenic *Escherichia coli* ST131. *J. Infect. Dis.* **2013**, *208*, 921-928.
- (39) Meanwell, M. A. Synopsis of some recent tactical application of bioisosteres in drug design. *J. Med. Chem.* **2011**, *54*, 2529-2591.
- (40) Prieto, M.; Zurita, E.; Rosa, E.; Luño, L.; Lloyd-Williams, P.; Giralt, E. Arylboronic acids and arylpinacolboronate esters in Suzuki coupling reactions involving indoles. Partner role swapping and heterocycle protection. *J. Org. Chem.* **2004**, *69*, 6812-6820.
- (41) Schulz, M. J.; Coats, S. J.; Hlasta, D. J. Microwave-assisted preparation of aryltetrazoleboronate esters. *Org. Lett.* **2004**, *6*, 3265-3268.
- (42) Devos, A.; Remion, J.; Frisque-Hesbain, A. M.; Colens, A.; Ghosez, L. Synthesis of acyl halides under very mild conditions. *J. Chem. Soc. Chem. Commun.* **1979**, 1180-1181.
- (43) Rabbani, S.; Jiang, X.; Schwardt, O.; Ernst, B. Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.* **2010**, *407*, 188-195.
- (44) Waetherman, R.V.; Kiessling, L.L. Fluorescence anisotropy assay reveals affinities of C- and O-glycosides for concanavalin A. *J. Org. Chem.* **1996**, *61*, 534-538.
- (45) Cer, R. Z.; Mudunuri, U.; Stephens, R.; Lebeda, F. J. IC50-to-Ki: a web-based tool for converting IC50 to Ki values for inhibitors of enzyme activity and ligand binding. *Nucleic Acids Research*, **2009**, *37*, W441-W445.
- (46) Lynch, B. A.; Loiacono, K. A.; Tiong, C. L.; Adams, S. E.; MacNeil, I. A. A Fluorescence Polarization Based Src-SH2 Binding Assay. *Anal. Biochem.* **1997**, *247*, 77-82.

-
- (47) Wu, P.; Brasseur, M.; Schindler, U. A High-Throughput STAT Binding Assay Using Fluorescence Polarization *Anal. Biochem.* **1997**, *249*, 29-36.
- (48) Huang, X. Fluorescence Polarization Competition Assay: The Range of Resolvable Inhibitor Potency Is Limited by the Affinity of the Fluorescent Ligand *J. Biomol. Screen.* **2003**, *8*, 34-38.
- (49) Cooper, A. *Biophysical Chemistry*, 2nd Edition. RSC Publishing. Cambridge UK. **2011**, pp. 122-123.
- (50) Scharenberg, M.; Jiang, X.; Pang, L.; Navarra, G.; Rabbani, S.; Binder, F.; Schwardt, O.; Ernst, B. Kinetic properties of carbohydrate-lectin interactions: FimH antagonists. *ChemMedChem* **2014**, *9*, 78-83.
- (51) Cabani, S.; Gianni, P.; Mollica, V.; Lepori, L. Group contribution to the thermodynamic properties of non-ionic solutes in dilute aqueous solution. *J. Solution Chem.* **1981**, *10*, 563-595.
- (52) Hansch, C.; Leo, A.; Taft, R.W. A survey of Hammett substituent constants and resonance and field parameters. *Chem. Rev.* **1991**, *91*, 165-195.
- (53) Chen, A.; Wadso, I. Simultaneous determination of ΔG , ΔH and ΔS by an automatic microcalorimetric titration technique: application to protein ligand binding. *J Biochem Biophys Meth*, **1982**, *6*, 307-316.
- (54) Freire, E.; Mayorga, O.L.; Straume, M. Isothermal titration calorimetry. *Anal Chem*, **1990**, *62*, 950A-959A.
- (55) Wiseman, T.; Williston, S.; Brandts, J. F.; Lin, L.-N. Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **1989**, *179*, 131-137.

- (56) Turnbull, W. B.; Daranas A.H. On the value of c : can low affinity systems be studied by isothermal titration calorimetry? *J. Am. Chem. Soc.* **2003**, *125*, 14859-14866.
- (57) Sigurskjold B. W. Exact analysis of competition ligand binding by displacement isothermal titration calorimetry. *Anal. Biochem.* **2000**, *277*, 260-266.
- (58) Velazquez-Campoy A.; Freire E. Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nat. Protoc.* **2006**, *1*, 186-191.
- (59) Dearden, J. C.; Bresnen, G. M. The measurement of partition coefficients. *QSAR Comb. Sci.* **1988**, *7*, 133-144.
- (60) Kansy, M.; Senner, F.; Gubernator, K. Physicochemical high throughput screening: parallel artificial membrane permeation assay in the description of passive absorption processes. *J. Med. Chem.* **1998**, *41*, 1007-1010.
- (61) Hubatsch, I.; Ragnarsson, E. G. E.; Artursson, P. Determination of drug permeability and prediction of drug absorption in Caco-2 monolayers. *Nat. Protoc.* **2007**, *2*, 2111-2119.
- (62) Banker, M. J.; Clark, T. H.; Williams, J. A. Development and validation of a 96-well equilibrium dialysis apparatus for measuring plasma protein binding. *J. Pharm. Sci.* **2003**, *92*, 967-974.
- (63) Obach, R. S. Prediction of human clearance of twenty-nine drugs from hepatic microsomal intrinsic clearance data: an examination of in vitro half-life approach and nonspecific binding to microsomes. *Drug Metab. Dispos.* **1999**, *27*, 1350-1359.
- (64) Chaturvedi, P. R.; Decker, C. J.; Odinecs, A. Prediction of pharmacokinetic properties using experimental approaches during early drug discovery. *Curr. Opin. Chem. Biol.* **2001**, *5*, 452-463.

- (65) Di, L.; Kerns, E. H. Profiling drug-like properties in discovery research. *Curr. Opin. Chem. Biol.* **2003**, *7*, 402-408.
- (66) Lipinski, C. A. Drug-like properties and the causes of poor solubility and poor permeability. *J. Pharmacol. Toxicol. Methods* **2000**, *44*, 235-249.
- (67) Curatolo, W. Physical chemical properties of oral drug candidates in the discovery and exploratory development settings. *Pharm. Sci. Techn. Today* **1998**, *1*, 387-393.
- (68) Ishikawa, M.; Hashimoto, Y. Improvement in aqueous solubility in small molecule drug discovery programs by disruption of molecular planarity and symmetry. *J. Med. Chem.* **2011**, *54*, 1539-1554.
- (69) Avdeef, A.; Bendels, S.; Di, L.; Faller, B.; Kansy, M.; Sugano, K.; Yamauchi, Y. PAMPA – critical factors for better predictions of absorption. *J. Pharm. Sci.* **2007**, *96*, 2893-2909.
- (70) Artursson, P.; Karlsson, J. Correlation between oral drug absorption in humans and apparent drug permeability coefficients in human intestinal epithelial (Caco-2) cells. *Biochem. Biophys. Res. Commun.* **1991**, *175*, 880-885.
- (71) Feng, B.; LaPerle, J. L.; Chang, G.; Varma, M. V. S. Renal clearance in drug discovery and development: molecular descriptors, drug transporters and disease state. *Expert Opin. Drug. Metab. Toxicol.* **2010**, *6*, 939-952.
- (72) Schmidt, S.; Gonzalez, D.; Derendorf, H. Significance of protein binding in pharmacokinetics and pharmacodynamics. *J. Pharm. Sci.* **2010**, *99*, 1107-1122.
- (73) Weisiger, R. A. Dissociation from albumin: A potentially rate-limiting step in the clearance of substances by the liver. *Proc. Natl. Acad. Sci. USA* **1985**, *82*, 1563-1567.
- (74) Smith, D. A.; Jones, B. C.; Walker, D. K. Design of drugs involving the concepts and theories of drug metabolism and pharmacokinetics. *Med. Res. Rev.* **1996**, *16*, 243-266.

- (75) Van de Waterbeemd, H.; Smith, D. A.; Beaumont, K.; Walker, D. K. Property-based design: optimization of drug absorption and pharmacokinetics. *J. Med. Chem.* **2001**, *44*, 1313-1333.
- (76) Varma, M. V. S.; Feng, B.; Obach, R. S.; Troutman, M. D.; Chupka, J.; Miller, H. R.; El-Kattan, A. Physicochemical determinants of human renal clearance. *J. Med. Chem.* **2009**, *52*, 4844-4852.
- (77) Waring, M. J. Lipophilicity in drug discovery. *Expert Opin. Drug Discov.* **2010**, *5*, 235-248.
- (78) Zhang Y.; Huo M.; Solver, P.K. An Add-in program for pharmacokinetic and pharmacodynamic data analysis in Microsoft Excel, *Computer Methods and Programs in Biomedicine* **2010**, *99*, 306-314.
- (79) Scharenberg, M.; Abgottspon, D.; Cicek, E.; Jiang, X.; Schwardt, O.; Rabbani, S.; Ernst, B. Cytometry-Based Assay for Screening FimH Antagonists. *Assay and Drug Development Technologies* **2011**, *9*, 455-464.
- (80) Hooton, T. M. Fluoroquinolones and resistance in the treatment of uncomplicated urinary tract infection. *Int. J. Antimicrob. Agents* **2003**, *22*, 65-72.
- (81) Jakobsen, L.; Cattoir, V.; Hammerum, A. M.; Nordmann, P.; Frimodt-Møller N. Impact of low-level fluoroquinolone resistance genes qnrA1, qnrB19, and qnrS1 on ciprofloxacin treatment of *Escherichia coli* urinary tract infection in murine model. Poster presented at *ICAAC 2010. 50th Interscience Conference on Antimicrobial Agents and Chemotherapy*, 2010, Sep 12-15, Boston, MA.
- (82) Mulvey, M. A. Adhesion and entry of uropathogenic *Escherichia coli*. *Cell Microbiol.* **2002**, *4*, 257– 271.

-
- (83) Ballatore, C.; Huryn, D. M.; Smith, A. B. Carboxylic acid (bio)isosteres in drug design. *ChemMedChem* **2013**, *8*, 385-395.
- (84) Houtman, J. C.; Brown, P. C.; Bowden, B.; Yamaguchi, H.; Appella, E.; Samelson, L. E.; Schuck, P. Studying multisite binary and ternary protein interactions by global analysis of isothermal titration calorimetry data in SEDPHAT: application to adaptor protein complexes in cell signaling. *Protein Sci.* **2007**, *16*, 30-42.
- (85) Maestro, version 9.3, Schrödinger, LLC, New York, NY, 2012.
- (86) MacroModel, version 9.9, Schrödinger, LLC, New York, NY, 2012.
- (87) AMSOL, version 7.1, Hawkins, G. D.; Giesen, D. J.; Lynch, G. C.; Chambers, C. C.; Rossi, I.; Storer, J. W.; Li, J.; Thompson, J. D.; Winget, P.; Lynch, B. J.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G. University of Minnesota, Minneapolis, 2003; based in part on AMPAC, version 2.1, Liotard, D. A.; Healy, E. F.; Ruiz, J. M.; Dewar, M. J. S.
- (88) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. Model for aqueous solvation based on class IV atomic charges and first solvation shell effects. *J. Phys. Chem.* **1996**, *100*, 16385-16398.
- (89) Dewar, M. J. S.; Zoebisch, E. G.; Healy, E. F.; Stewart, J. J. P. AM1 - A new general-purpose quantum-mechanical molecular-model (Vol. 107, PG 3902, 1985). *J. Am. Chem. Soc.* **1993**, *115*, 5348-5348.
- (90) Mulvey, M. A.; Schilling, J. D.; Hultgren, S. J. Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect. Immun.* **2001**, *69*, 4572-4579.

- (91) Kabsch, W. Automatic Processing of Rotation Diffraction Data from Crystals of Initially Unknown Symmetry and Cell Constants. *J. Appl. Crystallogr.* **1993**, 26, 795-800.
- (92) Leslie, A. G. W. The integration of macromolecular diffraction data. *Acta Crystallogr. D* **2006**, 62, 48-57.
- (93) Winn, M. D.; Ballard, C. C.; Cowtan, K. D.; Dodson, E. J.; Emsley, P.; Evans, P. R.; Keegan, R. M.; Krissinel, E. B.; Leslie, A. G. W.; McCoy, A.; McNicholas, S. J.; Murshudov, G. N.; Pannu, N. S.; Potterton, E. A.; Powell, H. R.; Read, R. J.; Vagin, A.; Wilson, K. S. Overview of the CCP4 suite and current developments. *Acta Crystallogr. D* **2011**, 67, 235-242.
- (94) McCoy, A. J. Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr. D* **2007**, 63, 32-41.
- (95) Emsley, P.; Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D* **2004**, 60, 2126-2132.
- (96) Adams, P. D.; Grosse-Kunstleve, R. W.; Hung, L.-W.; Ioerger, T. R.; McCoy, A. J.; Moriarty, N. W.; Read, R. J.; Sacchettini, J. C.; Sauter, N. K.; Terwilliger, T. C. PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **2002**, 58, 1948–1954.
- (97) van Aalten, D. M. F.; Bywater, R.; Findlay, J. B. C.; Hendlich, M.; Hooft, R. W. W.; Vriend, G. PRODRG, a program for generating molecular topologies and unique molecular descriptors from coordinates of small molecules. *J. Comput. Aided Mol. Des.* **1996**, 10, 255-262.
- (98) Chen, V. B.; Arendall, W. B.; Headd, J. J.; Keedy, D. A.; Immormino, R. M.; Kapral, G. J.; Murray, L. W.; Richardson, J. S.; Richardson, D. C. MolProbity: all-atom

- structure validation for macromolecular crystallography. *Acta Crystallogr. D* **2010**, *66*, 12-21.
- (99) Bezençon, J.; Wittwer, M. B.; Cutting, B.; Smiesko, M.; Wagner, B.; Kansy, M.; Ernst, B. pK_a determination by 1H NMR spectroscopy – an old methodology revisited. *J. Pharm. Biomed. Anal.* **2014**, *93*, 147-155.
- (100) (a) VCCLAB, Virtual Computational Chemistry Laboratory, 2005, <http://www.vcclab.org> (accessed August 14, 2012); (b) Tetko, I. V.; Gasteiger, J.; Todeschini, R.; Mauri, A.; Livingstone, D.; Ertl, P.; Palyulin V. A.; Radchenko, E. V.; Zefirov, N. S.; Makarenko, A. S.; Tanchuk, V. Y.; Prokopenko, V. V. Virtual computational chemistry laboratory – design and description. *J. Comput. Aided Mol. Des.* **2005**, *19*, 453-463.
- (101) Diehl K-H.; Hull R. A. Good Practice Guide to the Administration of Substances and Removal of Blood, Including Routes and Volumes. *J. Appl. Toxicol.* **2001**, *21*, 15-23.

Chapter 3.3. – Manuscript 5

The tyrosine gate of FimH – conformational analysis by NMR and X-ray

Preliminary manuscript formatted according to the requirements of *ChemBioChem*.

Contributions

- | | |
|---------------|---|
| B. Fiege | <ul style="list-style-type: none">• NMR measurements and data analysis.• Preparation of the manuscript. |
| R. C. Preston | <ul style="list-style-type: none">• Thermodynamic profiling by ITC and data analysis.• Protein-ligand co-crystallography, data acquisition and data analysis.• Preparation of the manuscript. |
| R. P. Jakob | <ul style="list-style-type: none">• Protein-ligand co-crystallography, data acquisition and data analysis.• Preparation of the manuscript. |
| P. Zihlmann | <ul style="list-style-type: none">• Thermodynamic profiling by ITC and data analysis. |
| S. Rabbani | <ul style="list-style-type: none">• Protein expression and purification. |
| O. Schwardt | <ul style="list-style-type: none">• Compound synthesis. |
| X. Jiang | <ul style="list-style-type: none">• Compound synthesis. |
| T. Maier | <ul style="list-style-type: none">• Supervision of protein-ligand co-crystallography.• Preparation of the manuscript. |
-

Understanding carbohydrate-lectin interaction

The tyrosine gate of FimH – conformational analysis by NMR and X-ray

Brigitte Fiege,^[a] Roland C. Preston,^[a] Roman P. Jakob,^[b] Pascal Zihlmann,^[a] Said Rabbani,^[a] Oliver Schwardt,^[a] Xiaohua Jiang,^[a] Timm Maier,^[b] and Beat Ernst*^[a]

Urinary tract infections caused by uropathogenic E. coli are among the most prevalent infectious diseases. The mannose-specific lectin FimH mediates the adhesion of the bacteria to the urothelium enabling host cell invasion and recurrent infections. An attractive alternative to antibiotics treatment is the development of FimH antagonists that mimic the physiological ligand. A large variety of candidate drugs have been developed and characterized using in vitro studies and animal models. Here we present the X-ray co-crystal structures of FimH with four antagonist classes for three of which structural data were unavailable to date. We used NMR spectroscopy

to further characterize FimH-antagonist interactions by chemical shift perturbation. The analysis allowed a clear determination of the conformation of the tyrosine gate motif that is crucial for interaction with aglycone moieties and was not obvious from X-ray structural data alone. Finally, ITC experiments provided insight into the thermodynamics of antagonist binding. In conjunction with the structural information from X-ray and NMR experiments the results suggest a mechanism for the often-observed enthalpy-entropy compensation of FimH antagonists that may play a role in fine-tuning of the interaction.

Introduction

Urinary tract infection (UTI) is one of the most frequent infectious disease affecting millions of people every year.^[1] Women have a 50% risk experiencing at least one symptomatic UTI during their lifetime. The large majority of UTIs is caused by uropathogenic *Escherichia coli* (UPEC) that are able to invade the urothelial cells in the bladder, form biofilms and cause recurrent infections.^[2] To date, UTIs are mainly treated with antibiotics leading to the ubiquitous problem of increasing antimicrobial resistance.^[3] Therefore, new treatment strategies are urgently needed.

Adherence of UPEC to the urothelial surface is mediated via the mannose-specific lectin FimH located at the tip of bacterial type 1 pili.^[4] FimH therefore represents a major virulence factor of UPEC. It consists of two immunoglobulin-like domains: the N-terminal lectin domain and – connected by a short linker – the C-terminal pilin domain. The lectin domain encloses the carbohydrate recognition domain (CRD). The pilin domain anchors the adhesin to the pilus and regulates the switch between the two conformational states of the CRD with high and low affinity for mannosides, respectively.^[5] Most *in vitro* studies on the interaction of FimH with FimH-antagonists have been performed with only the lectin domain FimH-CRD that is trapped in the high-affinity state.^[6] The cellular ligand for FimH is the mannosylated glycoprotein uroplakin Ia present on urothelial cells.^[7] Mannose-based FimH antagonists compete with the interaction and prevent the bacterial adhesion and hence the infection. A substantial advantage of this anti-adhesion therapy

over antibiotics treatment is the reduced risk of resistance development, since no direct selection pressure is imposed on the pathogen.^[8] The first successful demonstration of the anti-adhesion strategy was the protective effect of methyl α -D-mannoside in a UTI mouse model.^[9] Since then, great progress has been made in the optimization of FimH antagonists with affinities in the low nanomolar range.^[10] In part, these modifications are the result of rational drug design based on X-ray crystal structures of FimH bound to α -D-mannosides.^[10c, 11]

The mannose moiety is tightly bound and involved in an extended hydrogen bond network. A promising perspective for optimizing binding is offered by a rim of hydrophobic residues cresting the entrance to the mannose pocket. This so-called tyrosine gate, consisting of Tyr48 and Tyr137 with Ile52 positioned in between, is involved in the interaction with oligomannosides as revealed from the X-ray crystal structure of

[a] Dr. B. Fiege, Dr. R. C. Preston, P. Zihlmann, Dr. S. Rabbani, Dr. O. Schwardt, Dr. X. Jiang, Prof. B. Ernst
Institute of Molecular Pharmacy
University of Basel
Klingelbergstrasse 50, 4056 Basel (Switzerland)
Fax: (+41) 61 2671552
E-mail: beat.ernst@unibas.ch

[b] Dr. R. P. Jakob, Prof. T. Maier
Structural Biology, Biozentrum
University of Basel
Klingelbergstrasse 70, 4056 Basel (Switzerland)

Supporting information for this article is available on the WWW under <http://www.chembiochem.org> or from the author.

FimH-CRD with oligomannose-3.^[12] For drug design, the tyrosine gate was exploited for hydrophobic stacking interactions with suitable aglycones. In particular the finding that hydrophobic aryl and alkyl aglycones show stronger affinities led to many promising drug candidates. Through binding studies with a series of alkyl mannosides, *n*-heptyl α -D-mannoside (**1**) was identified as the most efficient non-aromatic FimH binder.^[10c] For aryl and heteroaryl derivatives a large diversity of antagonists differing in the number of aromatic rings and substituents, and the type of linkers exist. Chemically easy accessible aromatic derivatives such as 4-nitrophenyl and 4-methylumbelliferyl mannosides have been discovered quite early.^[10a, 10c] They were followed by derivatives of squaric acid,^[10d, 10j] biphenyls and other diaryls,^[10e, 10f, 10k, 10m] triazoles^[10g] and indolines^[10h] among others.

In our group, we routinely test new candidates for *in vitro* binding to FimH-CRD^[10f-j, 13] and to UPEC^[14] as well as for efficacy in a UTI mouse model.^[15] A special focus is placed on the PK/PD properties for oral bioavailability.^[10m] The ideal FimH antagonist for UTI treatment needs to be balanced between a reasonable solubility for effective dosage and a certain degree of lipophilicity for efficient membrane permeation during oral absorption. Despite the availability of structural information, rational design of FimH antagonists not always led to the expected improvement in binding affinities. Specifically, thermodynamic profiles from isothermal titration calorimetry (ITC) frequently revealed enthalpy-entropy compensations,^[10i, 10m] the reasons of which are not yet fully understood.

In this study, we analyzed the binding of a representative set of FimH antagonists to FimH-CRD through a combination of high-resolution X-ray crystal structures and NMR chemical shift perturbation (CSP) experiments. NMR spectroscopy allows the study of protein-carbohydrate interactions in solution, thus maintaining the molecule's natural dynamics behavior.^[16] In contrast, X-ray crystallography provides structural information from a "static" crystal with limited flexibility of the protein. Both methods are therefore highly complementary. The potential of NMR spectroscopy to contribute to the drug design process has been widely acknowledged.^[17] The chemical shifts of protein resonances are highly sensitive to the chemical environment e.g. to hydrogen bonds and aromatic ring currents.^[18] CSP effects hence identify residues in direct proximity to a bound ligand, and in addition indicate conformational changes in allosteric sites. We used ¹H-¹⁵N-HSQC experiments of ¹⁵N-enriched FimH-CRD in the presence of various antagonists to monitor the CSP effects on the backbone amide groups. Previous NMR studies of methyl α -D-mannoside^[6] and *n*-heptyl α -D-mannoside^[19] binding to FimH-CRD only noted a general match of the CSP maps and the ligand binding interface identified from X-ray co-crystal structures. With our analysis we demonstrate that NMR and X-ray structural data can be combined to characterize FimH-antagonist complex formation at unprecedented detail. In addition to the structural information, ITC experiments were performed to quantify the thermodynamics of antagonist binding to FimH-CRD.

Results and Discussion

FimH antagonist classes studied in this work.

In this study we examined FimH antagonists that are composed of an α -D-mannosyl residue linked to an aliphatic or aromatic aglycone (Table 1). *n*-Heptyl α -D-mannoside (**1**)^[10c] is the only

non-aromatic antagonists with a similar affinity (low nanomolar range) as the aromatic antagonists **2** to **7** (Table 1 and Table 2). Although its pharmacokinetic properties render **1** ineffective for oral administration, it is often used as a reference compound in screening studies. Biphenyl mannoside **2**^[10e] was rationally designed to make aromatic stacking interactions with the side chains of Tyr48 and Tyr137 in the tyrosine gate. Later on, the aryl moieties were derivatized to optimize this interaction as reported by our group and others.^[10e, 10i, 10k] An impressive example hereof is biphenyl mannoside **4** that carries an *ortho*-chlorine on the inner aromatic ring and a cyano group in the *para*-position of the outer ring as a bioorthogonal replacement of the carboxyl group present in antagonist **3**.^[10m] Both ring substituents of **4** reduce the electron density in the aglycone and thereby enhance the π - π interaction leading to a more than 10-fold affinity improvement in comparison to the unsubstituted antagonist **2** (Table 2). Besides the biphenyls **2** to **4**, three other compound classes with multiple aromatic rings in non-planar arrangements have been tested: **5** features a 5-nitroindoliny moiety *N*-linked to an inner phenyl ring;^[10h] **6** extends an *ortho*-chlorophenyl by squaric acid and *N*-methyl-piperazine;^[10j] and in **7** (see supporting information for synthesis) an amide bond is inserted between two phenyl rings with *ortho*-chlorine and *para*-carboxyl substituents.

In addition to high affinity antagonists, we studied methyl α -D-mannoside (**8**), *n*-butyl α -D-mannoside (**9**) and 1,5-anhydromannitol (**10**). Finally, the 4-deoxy-4-fluoro derivative **11** (see supporting information for synthesis) exhibits a dramatically reduced affinity compared to **1** emphasizing the importance of the hydrogen bond network in the mannose binding pocket.

X-ray crystal structures of FimH with antagonists.

We obtained high-resolution X-ray structures of FimH-CRD co-crystallized with antagonists **2** and **5-7** with a resolution between 1.00 and 1.65 Å (Table 1 and Table S1) and compared them to the known structures with antagonists **1** and **4**.^[10m] Unfortunately, biphenyl mannoside **3** could not be crystallized with FimH-CRD under the tested conditions. FimH-CRD comprises eleven mainly antiparallel β strands and two short helices (Figure 1). Superposition of the six protein structures reveals a high structural similarity with a backbone RMSD around 0.5 Å between any two structures. In all six structures the ligand is bound within the mannoside binding pocket. The mannosyl moieties form hydrogen bonds to the side chains of residues Asp54, Gln133, Asn135 and Asp140 and to the backbone of Phe1 and Asp47 (Figure S1). Tyr48 is involved in stacking interactions with the outer aromatic ring of the aglycones or in case of **1** with the alkyl chain of the ligands (Figure 1). Tyr137 stabilizes the loop containing Tyr48 by hydrophobic interaction with Ile52. In case of **6** and **7**, the side chain hydroxyl group of Tyr137 may form a hydrogen bond with the oxygen atoms of the squaric acid and the amide bond, respectively. Moreover, the backbone amides of Tyr137, Asn138 and Asp140 and the side chains of Asn135, Asn138 and Asp140 form a complex hydrogen bond network within the loop and with the 3- and 4-position of the mannosyl residue of the ligand.

Noteworthy, the tyrosine gate residues Tyr48 and Tyr137 nearly perfectly overlap in all six structures. The position of their side chains corresponds to the *closed conformation* of the tyrosine gate as described previously.^[10c, 10e, 10i, 20] The *closed conformation* was also observed in co-crystal structures of FimH-CRD with *n*-butyl α -D-mannoside (**9**) (PDB ID: 1UWF)^[10c] or with

a biphenyl mannoside with a methylcarboxylate group in *meta*-position of the outer aromatic ring (PDB ID: 3MCY).^[10e]

Table 1. FimH antagonists analyzed for binding to FimH-CRD using X-ray crystallography and/or CSP NMR experiments (indicated with +). For already published X-ray structures, the reference and PDB IDs are given.

Compd	Structure	X-ray	NMR
1		+ (4BUQ)	+
2		+ (ID)	+
3		n.d.	+
4		+ (4CST)	+
5		+ (ID)	+
6		+ (ID)	+
7		+ (ID)	+
8		n.d.	+
9		+ (1UWF, 1TR7) ^[10c]	+
10		n.d.	+
11		n.d.	+

A significantly different orientation of the tyrosine gate termed the *open conformation* was observed in the crystal structure of the full-length FimH in complex with α -D-mannose stabilized by the chaperone FimC (PDB ID: 1KLF),^[11] and of FimH-CRD with

1,2-ethanediol in the mannose binding site (PDB ID: 4AUU).^[20] The latter can be viewed as a “pseudo-apo” structure, although the bound 1,2-ethanediol could still have an influence on the binding site. While the Tyr137 side chain in these structures remains in the orientation present in the *closed conformation*, the Tyr48 side chain is rotated towards the mannose pocket by ca. 3.5 Å regarding the aromatic ring centers. This *open conformation* of FimH-CRD was also observed in the co-crystal structure with the native ligand epitope oligomannose-3, Man α 1,3(Man α 1,6)Man β 1,4GlcNAc β 1,4GlcNAc (PDB ID: 2VCO).^[12] Here, the α 1,3-linked mannose at the nonreducing end is recognized in the orientation typical for all mannosides. This brings the first GlcNAc residue into a position where it would clash with Tyr48 in the *closed conformation*. Considering the limited conformational freedom around the glycosidic linkages it can be speculated that oligomannose-3 can only bind to FimH in the *open conformation*.

In computational studies, FimH antagonists with flexible aglycones have been docked to FimH in both conformations of the tyrosine gate.^[10c, 10i] As an experimental verification, *n*-butyl α -D-mannoside (**9**)^[10c] and several monoaryl antagonists^[21] have been co-crystallized with FimH-CRD in both the *open* and *closed* form. Wellens et al. proposed that the *open conformation* represents the minimal energy conformer of FimH and the *closed conformation* is only stabilized by favorable interactions with hydrophobic aglycones, corresponding to a conformation selection process.^[20] Nevertheless, an induced fit mechanism also seems possible. To conclude, Tyr48 can be regarded as a key element in ligand interaction and the knowledge of the Tyr48 side chain conformation is critical to the discussion. Unfortunately, in many FimH-mannoside co-crystal structures the binding pocket residues or the ligand are involved in crystal contacts potentially affecting the binding pocket geometry.^[20-21] Therefore, more experiments are needed to study the structure and dynamics of FimH-antagonist complex formation in solution.

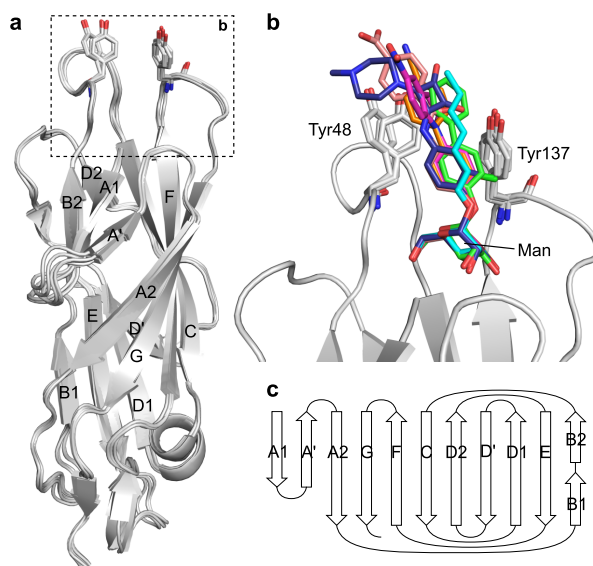


Figure 1. a), b) Superposition of X-ray co-crystal structures of FimH-CRD (grey, Tyr48 and Tyr137 shown as sticks) in complex with antagonists (shown as colored sticks in b) **1** (cyan), **2** (green), **4** (orange), **5** (magenta), **6** (blue) and **7** (salmon). c) Beta sheet topology diagram of FimH-CRD according to ^[10c, 22].

NMR experiments with FimH-CRD.

We performed NMR experiments to obtain structural information on FimH-antagonist complexes in solution. This allows a full retention of the molecule dynamics and avoids potential distortion from crystal packing. The binding of antagonists to ^{15}N -labeled FimH-CRD (173 residues) was monitored by CSP analysis of ^1H - ^{15}N -HSQC fingerprint spectra depicting the backbone amide groups. The protein contained a C-terminal 6His-tag for efficient affinity purification that does not influence mannoside binding.^[13] A complete backbone resonance assignment was performed by measuring a set of triple resonance NMR spectra with a ^{13}C , ^{15}N -labeled FimH-CRD sample. The high chemical shift dispersion of the amide resonances allowed sequential backbone assignment mainly on the basis of HNCACB and CBCACONH spectra.^[23] A total of 152 (94.4%) of 161 assignable residues could be assigned.^[24] The 6His-tag remained unassigned. Three residues were absent in the ^1H , ^{15}N -HSQC spectrum: the backbone and side chain signals of Asn96 and Arg98 which are located in a loop close to the binding pocket, as well as Gly79 located in an exposed loop ca. 20 Å away from the binding pocket. Carbon and proton signals of the side chains of these three residues were observed as (i-1) correlations of the succeeding residues

suggesting that fast exchange with the solvent and not (solely) conformational exchange was responsible for the signal absence. The assignment matches that from a recently published study of FimH-CRD (BMRB entry: 19066)^[19] that lacked a 6His-tag. The authors were able to observe and assign Gly79, Asn96 and Arg98 amide signals. This is likely due to different measuring conditions, in particular a lower pH of 6.0^[19] compared to pH 6.8 (for assignment) and pH 7.0 (for CSP experiments) in this study.

Chemical shift perturbation experiments. We subjected ^{15}N -labeled FimH-CRD to CSP experiments with antagonists **1** to **11** (Table 1). Methyl (**8**) and *n*-heptyl α -D-mannoside (**1**) have been analyzed in similar studies before.^[6, 19] With all compounds, separate signals for the bound and free form were observed in the ^1H , ^{15}N -HSQC spectra at 500 MHz indicating slow exchange on the NMR time scale. The observation is in accordance with nanomolar to low micromolar affinities (Table 2) and with kinetic data obtained from surface plasmon resonance experiments.^[25] Interestingly, even antagonist **11** that binds to FimH-CRD with a largely diminished affinity of 83 μM (Table 2) was found to be in slow or intermediate exchange. Slow exchange kinetics may be a hint to conformational rearrangements during complex formation.

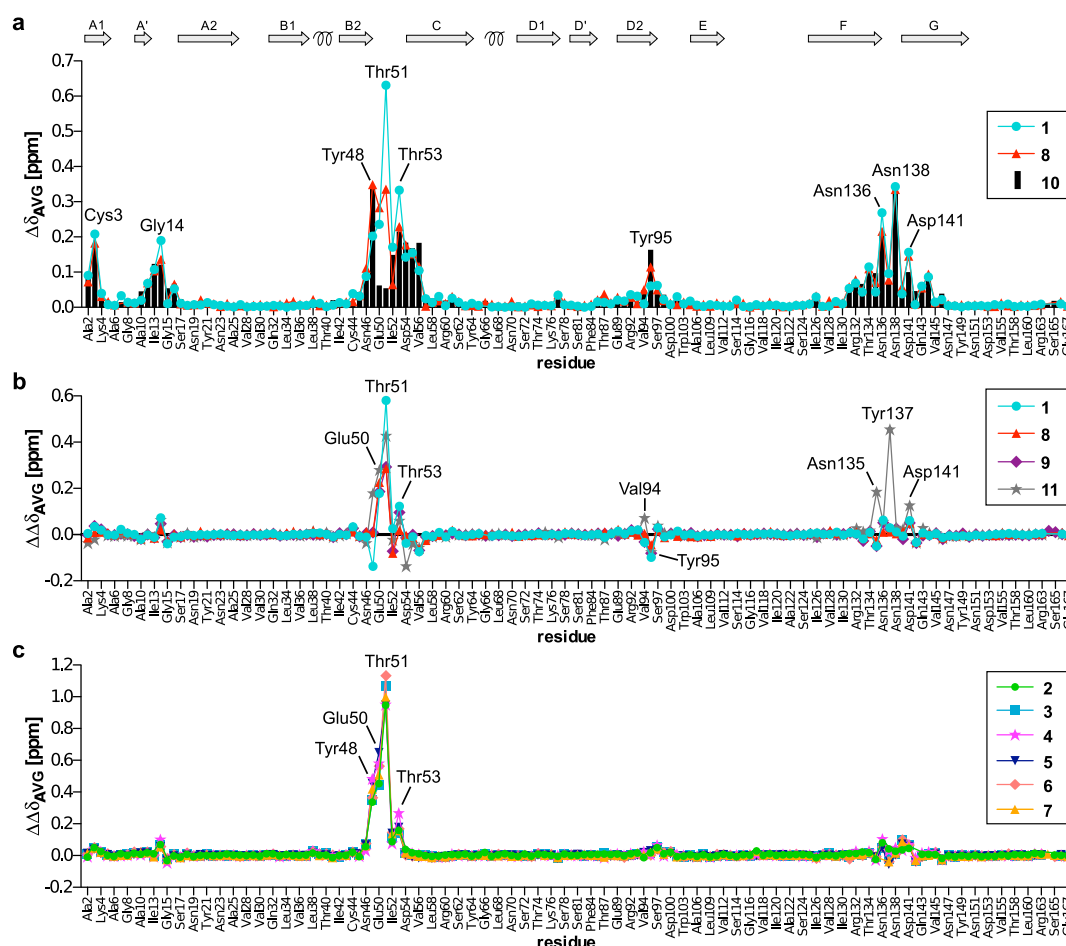


Figure 2. Chemical shift changes of FimH-CRD backbone amide signals upon addition of FimH antagonists. a) Absolute combined chemical shift changes $\Delta\delta_{\text{AVG}}$ of **1**, **8** and **10**. b), c) Differential CSP effects $\Delta\Delta\delta_{\text{AVG}}$ of aliphatic (b) and aromatic (c) ligands relative to 1,5-anhydromannitol (**10**). Secondary structure elements are schematized on top. Residues missing from the chart are Phe1 (N-terminus), 11 prolines (residues 12, 26, 49, 83, 85, 91, 102, 104, 111, 157 and 162), Gly79, Asn96 and Arg98 (not observed), Asp47, Tyr108 and Phe142 (overlap), Ser139 (exchange broadening), as well as the C-terminal 6His-tag (residues His168 to His173).

For the non-overlapping signals in the ^1H , ^{15}N -HSQC spectra, combined chemical shift changes, $\Delta\delta_{\text{AVG}}$, with respect to the protein signals in absence of ligand were determined (Figure 2a). 1,5-anhydromannitol (**10**) is the smallest structural motif that still shows specific binding to the mannose pocket of FimH-CRD. Its CSP effects were therefore used to calculate differential CSP effects, $\Delta\Delta\delta_{\text{AVG}}$, for the other ligands that should reflect mainly the influence of the aglycone moieties (Figure 2b,c). In presence of antagonists **1** to **11**, FimH-CRD displayed chemical shift changes for residues around the known mannose binding site. Only minor CSP effects below 0.1 ppm were observed for residues distal from this site. Since proton chemical shifts are highly sensitive to changes in hydrogen bond length,^[18] global conformational changes of FimH-CRD in response to antagonist binding can be excluded considering that the protein is comprised in large part of antiparallel beta sheets forming numerous hydrogen bonds.^[22] The observation is in agreement with the high structural similarity of the X-ray co-crystal structures discussed above. In presence of the “core” motif **10**, significant $\Delta\delta_{\text{AVG}}$ values above 0.1 ppm were observed for (in order of decreasing $\Delta\delta_{\text{AVG}}$) Tyr48, Asn138, Thr53, Asn136, Asp54, Val56, Cys3, Tyr55, Asp47, Tyr95, Ile52, Ile13, Gly14, Thr134 and Asn46. Most of these residues are directly constituting the binding pocket, in particular residues from the loop regions between beta sheets A' and A2 (Ile11 to Gly16), B2 and C (Asn46 to Val56) as well as F and G (Arg132 to Phe144). Of these, Tyr48, Ile52 and Tyr137 form the tyrosine gate that is known to interact with the aglycone moieties of FimH antagonists. In addition, CSP effects were observed for residues Ala2 and Cys3 from the bottom of the binding pocket, and for Tyr95 that forms a hydrogen bond with the pocket residue Asp54. For most residues, only very small $\Delta\Delta\delta_{\text{AVG}}$ values were observed. Larger deviations were observed for loop residues around Tyr137 in presence of **11** (see later discussion) and for loop residues around Tyr48.

CSP data report on the Tyr48 side chain conformation.

The strongest and most heterogeneous chemical shift changes of FimH-CRD with different antagonists were observed for the loop residues Tyr48, Glu50, Thr51, Ile52 and Thr53 (Figures 2 and 3). In the X-ray structures of FimH-CRD co-crystallized with **1**, **2** and **4** to **7**, the Tyr48 side chain directly interacts with the aglycone moieties and adopts a significantly different orientation compared to that in the “pseudo-apo” X-ray structure^[20] and in a recently solved NMR solution structure^[19] (Figure S2). Apart from Tyr48, the conformation of the loop residues is remarkably conserved in all FimH-CRD crystal structures. This includes hydrogen bonds between Tyr48 H^{N} and Asn46 $\text{O}^{\delta 1}$, between Glu50 H^{N} and its side chain oxygen (stabilized by interaction with the guanidinium group of Arg98, Figure S3) and between Thr51 H^{N} and Tyr48 O. We hypothesized that the observed backbone CSP effects of Tyr48 to Thr53 reflect mainly the conformational change of the Tyr48 side chain rather than direct ligand effects. In general, aromatic rings dramatically influence the chemical shifts of nearby nuclei through a local magnetic field induced by their delocalized π electrons. For protons located on the outside of the ring, the local magnetic field is parallel to the external field causing a downfield shift (higher ppm values). In contrast, protons located inside or above the ring experience an opposing field and hence an upfield shift. For Tyr48 and Glu50, we observed strong downfield shifts of both proton and ^{15}N chemical shifts in presence of antagonists while Thr51 experienced a strong proton upfield shift up to 1 ppm and above (Figure 3). The shifts were larger for aromatic compared to aliphatic aglycones. In the *open conformation* seen in the “pseudo-apo” crystal structure^[20] the Tyr48 ring is relatively far

away from the loop residues. In the *closed conformation*, it is rotated towards the loop such that Glu50 H^{N} comes closer and within 30° of the ring plane where it experiences a strong deshielding aromatic ring current. In contrast, Thr51 H^{N} is pointing nearly directly into the ring in a T-shaped orientation where a strong shielding field causes upfield shifts.^[26] Previous studies demonstrated that T-shaped N-H/ π interactions can significantly contribute to ligand binding affinities.^[27] Enhancement of the aromatic ring current effect by π - π stacking provides the explanation for the generally larger CSP effects observed for aromatic compared to aliphatic antagonists. Importantly, Glu50 and Thr51 show nearly no shifts in presence of **10** strongly indicating that FimH-CRD remains in the unperturbed *open conformation*.

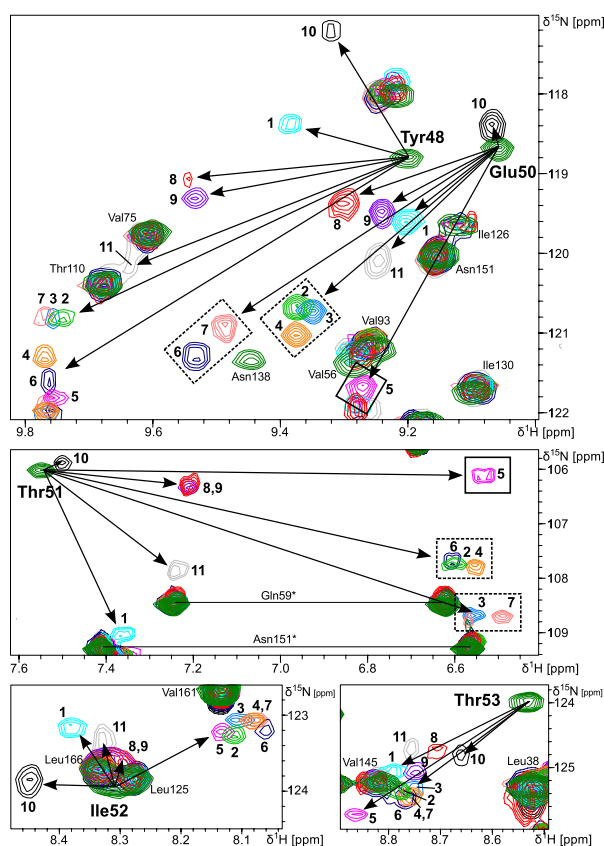


Figure 3. Chemical shift changes of the binding loop residues Tyr48, Glu50, Thr51, Ile52 and Thr53 in presence of antagonists **1** to **11**. Spectra in absence of ligand are colored in green. Signals corresponding to Tyr48 conformations a and b in Figure 4 are marked by dashed and solid line boxes, respectively.

A more detailed analysis of the CSP effects with aromatic antagonists allows even further differentiation of the Tyr48 orientation in the *closed conformation*. The signals of Glu50 and Thr51 group together for compounds **2**, **3**, **4**, **6** and **7** (Figure 3, dashed line boxes). With antagonist **5**, slightly different shifts were observed suggesting a different orientation of Tyr48. Aliphatic compounds **1** and **8-11** cannot be directly compared to the aromatic antagonists, since the aromatic ring current effect of Tyr48 as the main source of the shifts is drastically different. In good agreement with the NMR results, the high-resolution X-ray

structure of FimH-CRD in complex with **5** shows a distinct orientation of the Tyr48 ring that is tilted by ca. 40° compared to the co-crystal structures with compounds **1**, **2**, **6** and **7** (Figure 4, bottom). This tilt is likely the result of the unique geometry of the nonplanar indolinyphenyl moiety of **5** and the propensity of the system to optimize the stacking interaction. Interestingly, in the co-crystal structure with biphenyl derivative **4** the electron density clearly allows positioning of the Tyr48 ring in both conformations (Figure 4). We cannot find a confirmation for this in the NMR experiments, since only a single set of bound signals is observed with **3** and the CSP effects are very similar to that with compounds **2** and **6** with a single Tyr48 orientation in the crystal structure. Nevertheless it is conceivable that Tyr48 can retain a certain degree of flexibility in the bound state (see later discussion on ITC data). NMR chemical shifts of the binding loop residues thereby would show an average of fast exchanging Tyr48 orientations in the complex.

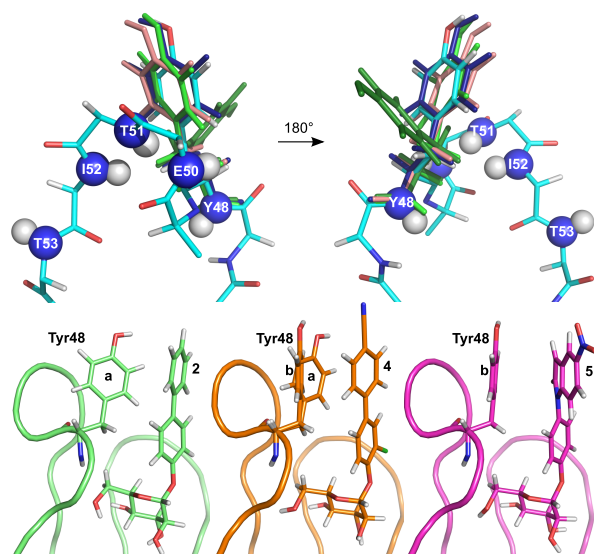


Figure 4. Orientation of the Tyr48 side chain in FimH-CRD crystal structures. Top: Loop residues Tyr48 to Thr53 in the co-crystal structure with **1** (backbone in cyan, amides shown as spheres); Tyr48 side chain in the “pseudo-apo” structure^[20] (green) and with **1** (cyan), **2** (pale green), **6** (dark blue) and **7** (salmon). Bottom: Different Tyr48 orientations (a and b) with **2**, **4** and **5**.

Water coordination in the FimH binding pocket. Some CSP effects are expected to be due to direct interactions of the protein with the mannosyl moiety. X-ray structural data implicate the formation of a direct hydrogen bond of Asp47 H^N to Man OH6 and of a water-mediated hydrogen bond of Gly14 H^N to Man OH2 (Figure 5). Against the expectation of a ¹H downfield shift from deshielding upon hydrogen bond formation, both amide protons showed an upfield shift (Figure 5). This rather points to a weakening of an existing hydrogen bond. Although other more complex effects cannot be excluded, the results indicate that Gly14 and Asp47 coordinate water in absence of ligand that is replaced upon mannoside binding by the ligand’s polar groups in an effectively weaker hydrogen bond. The higher degree of freedom of water molecules may indeed allow hydrogen bonds with more ideal geometries compared to the ligand. In the X-ray co-crystal structures with antagonists **1**, **2** and **4** to **7** the distance of the bridging water to Gly14 H^N is 0.1 to 0.2 Å closer than in the

“pseudo-apo” crystal structure in agreement with our hypothesis (Figure 5). Finally, molecular dynamics simulations confirmed the presence of a structural water molecule close to Gly14.^[28]

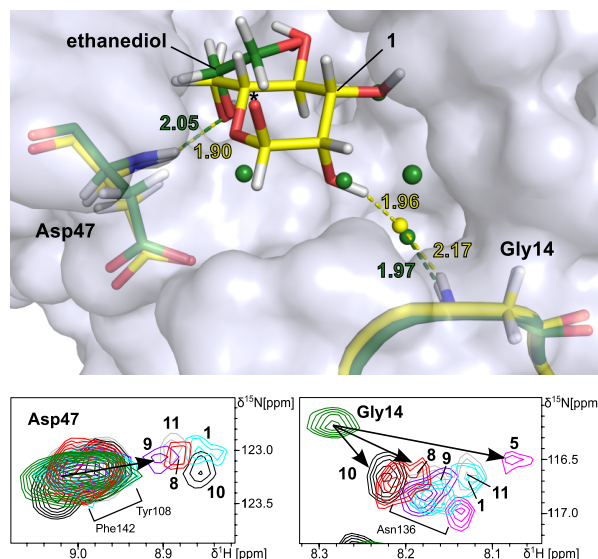


Figure 5. Water coordination in the FimH binding pocket. Top: X-ray structures of “pseudo-apo” FimH-CRD (green sticks and transparent surface)^[20] and in complex with **1** (yellow, asterisk indicates attachment point of aglycone); water molecules are shown as spheres; hydrogen bond lengths are given in Å. Bottom: Chemical shift changes of Asp47 (only assigned for non-aromatic antagonists) and Gly14 (only subset of spectra shown).

Hydrogen bond network in the binding loop with Tyr137.

The loop residues Asn135 to Asp141 display similar CSP effects in presence of all tested antagonists except **11** (Figure 2). X-ray co-crystal structures revealed a complex interresidual hydrogen bond network between backbone and side chain atoms within this loop (Figure 6). For example, Asn138 H^N forms a hydrogen bond to the side chain oxygen of Asn135. The strong deshielding of Asn138 H^N in presence of ligand suggests a strengthening of this hydrogen bond (Figure 6). The shift could also result from a conformational change of the aromatic ring of Tyr137. However, Tyr137 H^N displayed no or only very small shifts suggesting that its aromatic ring remains largely unchanged in agreement with very similar orientations in the X-ray structures. Furthermore, deshielding of Asp140 H^N with all ligands except **1** and **9** indicate a slight shortening of the hydrogen bond to the side chain oxygen of Asn138 (Figure 6). In summary, mannoside coordination seems to induce subtle but specific conformational changes in the hydrogen bond network within the binding loop. The differences in the CSP effects of Asn135, Tyr137 and Asp141 with the 4-deoxy-4-fluoro derivative **11** are hence caused by a different electronic environment of the fluorine atom that also compels a different loop arrangement through disruption of hydrogen bonds. The X-ray structural data are not entirely conclusive on the effect of mannoside binding to the loop conformation. In all six co-crystal structures discussed above, the loop residues are well ordered, but hydrogen bonds in the loop are identical within error to that in the “pseudo-apo” structure. Apparently, very small conformational variations notably in an exposed loop cannot be resolved in the X-ray structures despite the relatively high resolution.

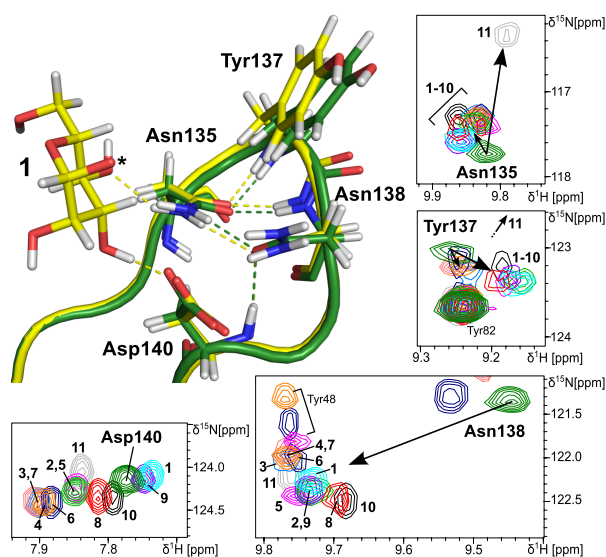


Figure 6. Hydrogen bond network in the binding loop with Tyr137. Top left: X-ray crystal structures of “pseudo-apo” FimH-CRD (green)^[20] and in complex with **1** (yellow, attachment point of aglycone marked with an asterisk); observed hydrogen bonds: Asn135 H^{δ22}–Asn138 O^{δ1}, Tyr137 H^N–Asn135 O^{δ1}, Asn138 H^N–Asn135 O^{δ1}, Asp140 H^N–Asn138 O^{δ1}; with ligand: Man OH3–Asp140 O^{δ1}, Asn135 H^{δ21}–Man O4. Bottom and right: Chemical shift changes of Asn138 and Asp140 with antagonists **1** to **11**.

Conformational changes in the pocket zipper motif. The CSP effects for Cys3 are unlikely to be caused by direct ligand effects considering its remote distance (>8.5 Å) from the bound mannose. In FimH-CRD, Cys3 is part of a highly structured and stable hydrogen bond network to neighboring beta strands further stabilized by a disulfide bond to Cys44 (Figure 7). This so-called “pocket zipper” motif is one of four regions that are suspected to play a key role in the conformational transition between the low and the high affinity states of in the full-length FimH.^[6, 29] Upon transition to the high affinity state, the “pocket zipper” becomes tightly hydrogen bonded causing a constriction of the binding pocket.^[6] The isolated lectin domain FimH-CRD is supposed to be locked in the high-affinity state. In the CSP experiments with FimH antagonists, we observed relatively uniform downfield shifts for Cys3 H^N that forms a hydrogen bond to the backbone oxygen of Ile11 (Figure 7). A similar albeit smaller effect was seen for Ile11 H^N that forms an adjacent hydrogen bond to the Cys3 backbone oxygen. This suggests a strengthening of the hydrogen bonds, although a contribution of nearby aromatic rings, in particular of Phe144 (ca. 3 Å from Cys H^N) seems also possible. In the six FimH-CRD crystal structures discussed above the hydrogen bond Cys3 H^N–Ile11 O is 0.2 to 0.3 Å shorter than in the “pseudo-apo” structures in agreement with the CSP effects (Figure 7). It is again emphasized that such subtle differences are difficult to substantiate even from high-resolution X-ray structures. A strengthening of the “pocket zipper” upon ligand binding could point to a residual conformational transition of FimH-CRD. However, we observed no CSP effects in the other allosteric regions confirming that the isolated lectin domain is not undergoing significant conformational adaptations characteristic for the full-length protein. Corresponding NMR solution studies of full-length FimH would be highly desirable but have so far been hampered by limited protein yield and stability.

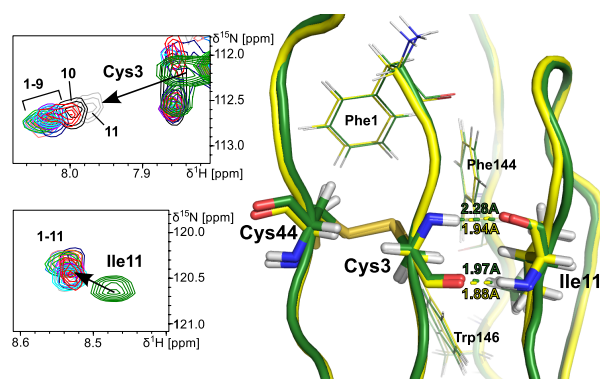


Figure 7. Conformational changes in the pocket zipper motif. Left: Chemical shift changes of Cys3 and Ile11 in presence of **1** to **11**. Right: X-ray crystal structures of “pseudo-apo” FimH-CRD (green)^[20] and in complex with **1** (yellow); hydrogen bond lengths are given in Å; aromatic residues in the vicinity are shown as lines.

Thermodynamics data from ITC experiments

FimH antagonists **1** to **11** have been subjected to ITC analysis with FimH-CRD (Table 2 and Table S2). To obtain reliable thermodynamics data, the high affinity of some of the compounds required the establishment of a competitive ITC assay^[30] (see supporting information for details). The results indicate that all tested antagonists bind enthalpy-driven with mostly unfavorable entropic contributions in accordance with previous studies.^[10i, 10m, 20] The enthalpy term $\Delta H^{\circ}_{\text{obs}} = -42.9$ kJ/mol of 1,5-anhydromannitol (**10**) contains the binding energies from van der Waals contacts and ten specific hydrogen bonds with the protein.^[10c, 11] These favorable terms are partly compensated by enthalpic penalties from desolvation of ligand and protein. The enthalpic cost for desolvation of a single hydroxyl group had been estimated at 29 kJ/mol.^[31] On the protein side, X-ray and NMR data and molecular dynamics simulations found evidence for highly structured hydrogen-bonded water in the binding pocket that is replaced by the hydroxyl groups of the ligand. The classic hydrophobic effect predicts a strong entropic gain for the release of bound water upon complex formation.^[32] However other studies also suggested an enthalpy-driven “non-classical” hydrophobic effect that is usually explained by water being forced into an enthalpically unfavorable configuration in the binding pocket.^[33] The entropy term of **10** suggest that any entropic gain from the release of water into the bulk is overcompensated by unfavorable contributions such as the loss of rotational and translational entropy of the ligand (estimated at ca. 25 kJ/mol in aqueous solutions)^[34] as well as conformational restriction of ligand and protein in the complex. Relevant for the latter is the formation of several hydrogen bonds that reduce the flexibility of the ligand's hydroxyl groups and the involved protein residues. Quantification of the individual entropic contributions is cumbersome and in fact still represents a major obstacle in current research.^[33, 35]

The ITC data for methyl α -D-mannoside (**8**) reveal a less favorable enthalpy ($\Delta\Delta H^{\circ}_{\text{obs}}: 5.7$ kJ/mol) and a smaller entropy penalty ($-T\Delta\Delta S^{\circ}_{\text{obs}}: -5.5$ kJ/mol) compared to **10**. The additional methoxy group of **8** is supposed to increase the binding enthalpy through van der Waals interactions with the protein (Ile13, Asp47 and Tyr48 within 4 Å). The effective enthalpy loss is likely due to an additional desolvation penalty for the methoxy group, while the entropy gain reflects the release of additional water molecules

into the bulk. From NMR CSP experiments we also expect a change in the Tyr48 side chain conformation for all mannosides with methyl or larger aglycones. The implications for the thermodynamics depend on whether the complex formation follows an induced fit mechanism or conformational selection or a mixture of both. In a recent solution NMR study of FimH-CRD, ^{15}N relaxation and CPMG relaxation dispersion experiments of the backbone amides did not provide any indication for multiple conformations of Tyr48 in the apo protein neither for a significant change in the backbone flexibility in presence of **1**.^[19] The data therefore rather argue for an induced-fit mechanism of the tyrosine gate instead of conformational selection. Nevertheless, conformational equilibria of Tyr48 in the apo state may exist in a time scale not easily accessible by NMR experiments (μs to ms). Elongation of the aglycone from methyl (**8**) to butyl (**9**) and heptyl (**1**) led to significant improvement of the binding enthalpy due to the enhanced C-H/ π interaction of the elongated alkyl chain with the Tyr48 ring. An expected improvement of the entropy term from the release of more water into the bulk is observed for **9** but not for **1** that even shows a loss of entropy compared to **8**. We supposed that the stacking interaction of the heptyl chain of **1** with Tyr48 leads to rigidification of both interaction partners and hence to significant entropic costs. For compound **11**, the 4-deoxy-4-fluoro analog of **1**, a significantly smaller enthalpy term was observed as expected from the disruption of the hydrogen bonds in the 4-position. The CSP experiments also indicated structural changes in the binding loop containing Tyr137 that can translate into enthalpic costs. Interestingly, the entropy is improved by more than 9 kJ/mol in comparison to **1** likely due to the higher flexibility of the protein binding pocket and the ligand.

For aromatic antagonists **2** to **7**, strongly favorable enthalpy terms ranging from -45.0 to -71.8 kJ/mol were obtained. Again, an increase in enthalpy is always accompanied by an increase of entropic costs as becomes obvious in the enthalpy-entropy compensation plot (Figure S4). X-ray structure data indicate that the outer aromatic rings of antagonists **2** and **4-7** make parallel-displaced π - π interactions with the Tyr48 ring in the *closed conformation* in analogy to the C-H/ π interaction of the heptyl chain of **1**. However, a direct comparison of the ITC data of aliphatic and aromatic antagonists is difficult in view of the large structural difference of the aglycones and because of the many contributing factors. The unsubstituted biphenyl mannoside **2** displayed the smallest enthalpy term and likewise the smallest entropic penalty of all aromatic antagonists. For the substituted biphenyls **3** and **4** we observed a significant increase of the enthalpy by 11.2 and 15.9 kJ/mol, respectively, in comparison to **2**. First, the *ortho*-chloro substituent is assumed to enhance the binding enthalpy by favorable van der Waals interactions with the protein in a small pocket formed by Ile52, Tyr137 and Asn138.^[10] Second, the electron-withdrawing character of the substituents allows a stronger π - π interaction of the aglycones with the Tyr48 ring. As already suggested for the comparison of **1** with **8**, a stronger stacking interaction likely accounts for a reduced flexibility in both the Tyr48 side chain and the aglycone leading to compensating entropic penalties. For antagonist **5**, the enthalpy is slightly better than for the best biphenyl compound **4**. The entropy compensation is only partial resulting in the strongest dissociation constant of all tested antagonists. Compounds **6** and **7** display the most favorable enthalpies but also the highest entropic penalties. **6** could suffer from entropic costs for rigidification of additional rotatable bonds in the aglycone moiety. **7** is structurally very similar to biphenyl **3** apart from insertion of an amid bond linkage. The dramatically enhanced enthalpy ($\Delta\Delta H^\circ_{\text{obs}}$:

-15.6 kJ/mol) and entropic costs ($-\Delta\Delta S^\circ_{\text{obs}}$: 17.0 kJ/mol) of **7** compared to **3** can only be explained by a very efficient stacking interaction with Tyr48 and subsequent loss of conformational entropy in the complex. From comparison of all ITC data it becomes clear, that the antagonist with the highest enthalpic contribution presumably from the strongest stacking interaction with Tyr48 is not necessarily the strongest binder. Antagonist **5** exhibits the best thermodynamic profile, although its enthalpy term is nearly 10 kJ/mol weaker than that of **7**. From the NMR and X-ray data we deduced a distinct orientation of Tyr48 in presence of **5** likely as a result of the unique geometry of its indolylphenyl aglycone. We assume that this complex conformation is optimal in terms of interaction efficiency and minimized compensating entropic penalties. However, with regard to the pharmacokinetic profiles of FimH-antagonists, biphenyl mannoside **4** was identified as the optimal lead compound for oral treatment of UTI as described in a separate publication.^[10m]

Table 2. Thermodynamics of FimH antagonists binding analyzed by ITC.

Compd	K_D [nM]	$\Delta G^\circ_{\text{obs}}$ [kJ/mol]	$\Delta H^\circ_{\text{obs}}$ [kJ/mol]	$-\Delta S^\circ_{\text{obs}}$ [kJ/mol]
1 ^[a]	28.9 (25.8 – 32.3)	-43.0	-50.3 (-50.2 – -50.7)	7.3
2 ^[a]	17.7 (14.1 – 22.3)	-44.2	-45.0 (-44.5 – -45.6)	0.8
3 ^[b]	3.5 (2.7 – 4.6)	-48.3	-56.2 (-55.9 – -56.6)	8.0
4 ^[b]	1.3 (1.1 – 1.5)	-50.7	-60.9 (-60.5 – -61.3)	10.1
5 ^[b]	1.0 (0.7 – 1.4)	-51.3	-62.1 (-61.5 – -62.7)	10.8
6 ^[a]	14.0 (12.7 – 15.4)	-44.8	-63.6 (-63.2 – -64.0)	18.8
7 ^[a]	6.2 (5.1 – 7.6)	-46.8	-71.8 (-71.0 – -72.7)	25.0
8 ^[a]	1222 (1064 – 1403)	-33.8	-37.2 (-36.4 – -38.0)	3.5
9 ^[a]	144 (127 – 163)	-39.1	-39.1 (-38.6 – -39.7)	0.1
10 ^[a]	1125 (980 – 1292)	-34.0	-42.9 (-42.1 – -43.8)	9.0
11 ^[a]	89990 (86140 – 94020)	-23.1	-21.1 (-20.8 – -21.6)	-2.0
[a] Direct ITC assay format.				
[b] Competitive ITC assay format using a medium-affinity ligand.				

Conclusion

We have combined NMR chemical shift perturbation (CSP) experiments with high-resolution X-ray structures to elucidate the interaction of the FimH lectin domain with antagonists in great detail. One advantage of NMR spectroscopy is the ability to identify even subtle conformational changes of the protein in solution. This is particularly helpful for the determination of ligand binding modes in cases where crystallization was unsuccessful or crystal packing effects may influence the protein and ligand conformation. We demonstrated that the CSP effects of FimH-CRD can be used as an indicator for the conformation of the tyrosine gate motif in the binding pocket. The *open conformation* was observed in the apo protein and with 1,5-anhydromannitol (**10**), while antagonists with aliphatic or aromatic aglycones bound the *closed conformation* in agreement with X-ray structural data.

The CSP data even revealed slight differences in the Tyr48 conformation as a result of different aglycone geometries. Furthermore, the NMR results report on the presence of highly structured water in the binding pocket of FimH-CRD. In combination with information from crystal structures and molecular modeling, CSP experiments therefore may help to analyze water coordination as an important contributor to the thermodynamics of ligand binding.

We also performed ITC experiments to access full thermodynamic profiles of the antagonists. The results suggest enthalpy-entropy compensation in which enthalpically favorable stacking of hydrophobic aglycones with the Tyr48 side chain leads to conformational restriction of both the protein and the ligand and hence to unfavorable entropy. NMR relaxation experiments of the backbone amide signals of FimH-CRD did not show any changes in the backbone flexibility upon antagonists binding. However, we suppose that flexibility of the protein side chains, in particular of Tyr48, could play a pivotal role in the modulation of the thermodynamics of FimH-antagonist binding. These effects can be dissected by NMR relaxation experiments for side chain dynamics and incorporation of the results into rational design may lead to further improved FimH antagonists urgently needed for the therapy of UTI.

Experimental Section

Protein preparation: FimH-CRD from the *E. coli* K-12 strain was expressed with a C-terminal thrombin cleavage site and a 6His-tag (FimH-CRD-Th-6His, 173 residues) following a previously published protocol.^[13] The clone containing the FimH-CRD construct was expressed in protease-deficient *E. coli* HM 125 at 30°C and 180 rpm in M9 minimal medium supplemented with 100 µg/mL ampicillin. The protein expression was induced by 1 mM IPTG at an OD₆₀₀ of 0.8. The cells were further cultivated for 16 h, harvested by centrifugation for 20 min at 5000 rpm and 4°C. The pellet was resuspended in lysis buffer containing 50 mM Tris pH 7.4, 150 mM NaCl, 5 mM EDTA and 1 mg/mL polymyxin B sulfate. The supernatant containing the periplasmic extract was dialyzed against sodium phosphate buffer and purified on Ni-NTA columns. For crystallization, the protein was dialyzed against 20 mM HEPES buffer pH 7.4. For production of ¹⁵N- and ¹³C, ¹⁵N-labeled FimH-CRD-Th-6His for NMR experiments, *E. coli* HM 125 was cultivated in M9 minimal medium containing 1 g/L ¹⁵NH₄Cl or 1 g/L ¹⁵NH₄Cl and 2 g/L ¹³C-glucose (Sigma, Switzerland) as the only sources of nitrogen and carbon, respectively. The labeled proteins were purified as described above and dialyzed against 25 mM phosphate buffer pH 6.8 (¹³C, ¹⁵N-FimH-CRD) or 20 mM phosphate buffer pH 7.0 (¹⁵N-FimH-CRD). The purity of the proteins was verified by non-reducing SDS-PAGE analysis, and the concentration was determined by UV absorption (NanoDrop ND-1000, Thermo Scientific). The molecular weight was determined by mass spectrometry for ¹⁵N-FimH-CRD (18860.2 Da) and ¹³C, ¹⁵N-FimH-CRD (19687.0 Da) demonstrating an isotope incorporation of > 99.9%.

NMR spectroscopy: NMR assignment experiments were measured at 298 K on a Bruker Avance III 700 MHz NMR spectrometer equipped with a 5 mm QCI-P cryogenic probe. A set of triple resonance NMR spectra of a 600 µM sample of ¹³C, ¹⁵N-FimH-CRD in 25 mM phosphate buffer pH 6.8, 90%/10% H₂O/D₂O in a 5 mm Shigemi Tube (Shigemi Inc., USA) was recorded for backbone assignment: HNCA, HN(CO)CA, HNCO, HN(CA)CO, HNCACB and CBCA(CO)NH. Protein stability during data acquisition was tested via regular inspection of 1D ¹H and ¹H, ¹⁵N-HSQC NMR spectra. Spectra were acquired and processed with Topspin 3.2 (Bruker BioSpin, Switzerland). CcpNmr Analysis (versions 2.2 and 2.3) was used for NMR resonance assignment.^[36] The backbone assignment of FimH-

CRD has been deposited in the Biological Magnetic Resonance Data Bank (<http://www.bmrb.wisc.edu>).

Chemical shift perturbation (CSP) experiments with FimH antagonists were performed with 100 to 200 µM ¹⁵N-FimH-CRD in 20 mM non-deuterated phosphate buffer pH 7.0. FimH antagonists were solved in D₂O at 2.5 to 100 mM stock concentrations. Biphenyl compound **4** was solved at 6.3 mM in D₂O with 40% of DMSO-*d*₆ due to poor solubility. The antagonists were added at slight molar excesses to ensure complete saturation of the protein. NMR samples were prepared in 3 mm NMR tubes (Hilgenberg, Germany) with 5% D₂O and *d*₄-TSP ((3-(trimethylsilyl)-2,2',3,3'-tetra-deuteriopropionic acid) added as internal reference. ¹H, ¹⁵N-HSQC spectra were measured on a Bruker Avance III 500 MHz spectrometer equipped with a BBO double resonance probe at 298 K. Spectra were acquired and processed with Topspin 2.1 and analyzed with CcpNmr Analysis (versions 2.2 and 2.3). All antagonists bound in the slow exchange regime and amide signals of the bound protein were assigned from chemical shift proximity and matching peak intensity. Combined chemical shift changes of FimH-CRD signals were calculated as weighted average of ¹H and ¹⁵N chemical shift changes according to the equation $\Delta\delta_{AV} = \sqrt{(\Delta\delta^1H)^2 + (0.2\Delta\delta^{15}N)^2}$.^[37]

FimH-antagonists co-crystallization: For crystallization trials, FimH-CRD (residues 1-158) at a final concentration of 18 mg/mL (ca. 1 mM) with a fivefold molar excess of ligand (5 mM) in 20 mM HEPES buffer pH 7.4 was used. Crystals were grown in sitting-drop vapor diffusion at 20°C as described before.^[10m] Precipitation solutions were determined by PEG/ION HT screen (Hampton Research, USA) and details for every crystal are given in the supporting information. In some cases streak seeding was performed with crystal seeds of FimH-CRD and compound **5** grown after one week. Crystals appeared within 24 h or in case of compound **6** after 4 weeks. Crystals were flash-cooled to 100 K with perfluoropolyether cryo oil (Hampton Research, USA). Data was collected with synchrotron radiation at the PXIII beamline of the Swiss Light Source (Paul Scherrer Institute, Switzerland).

Structure determination and refining: Data was indexed, integrated and scaled with XDS.^[38] Structures were solved by molecular replacement with PHASER^[39] using the FimH-CRD-*n*-butyl α -D-mannopyranoside complex (PDB ID: 1UWF^[10c]) as search model. The structures were built using the COOT software^[40] and periodically refined with the PHENIX software.^[41] Geometric restraints for the ligands were generated with PRODRG.^[42] Molprobity^[43] was used to validate the available atomic coordinates and to add protons for distance calculation. The structures were submitted to the Protein Data Bank with PDB codes 4X50 (**2**), 4X5Q (**5**), 4X5R (**6**) and 4X5P (**7**).

Isothermal titration calorimetry (ITC): All ITC experiments were performed with FimH-CRD-Th-6His using a MicroCal VP-ITC instrument (Malvern Instruments Ltd, Worcestershire, UK) with a sample cell volume of 1.4523 mL. Measurements were performed at 25°C with a stirring speed of 307 rpm and 10 µcal/s reference power. The protein was dialyzed against assay buffer (10 mM HEPES, 150 mM NaCl, pH 7.4). Injections of 3-12 µL ligand solution were added in 10 min intervals to the sample cell containing 5-50 µM protein. The *c*-values ($c = M_i(0) K_d^{-1}$, where *M_i(0)* is the initial protein concentration) were in a reliable range within 5 and 1'000 for compounds **1**, **2** and **6-10**. For compounds **3**, **4**, and **5** the *c*-values of the direct titrations were above 1'000 and therefore additional competitive ITC experiments were performed.^[30] The ligands were titrated into protein preincubated with an 8 to 9-fold excess of compound **12** (structure and synthesis in supporting information) resulting in sigmoidal titration curves. For the low affinity compound **11** the *c*-value was only 0.5, but reliable thermodynamics data could be extracted by fixing the stoichiometry to 1.0. Baseline correction and peak integration was performed using the Origin 7 software (OriginLab, USA). Baseline subtraction and curve fitting with the three variables *N* (concentration correction factor), *K_d* (dissociation constant), and ΔH° (change in enthalpy) were performed with

SEDPHAT version 10.40 (National Institute of Health).^[44] Global fitting to obtain K_D values was performed for the competitive titrations of compound **12** with **3**, **4** and **5** and for the direct titration of **12**. ΔH° and N were then obtained from fitting of the direct titrations of **3**, **4** and **5**. For directly titrated compounds **1**, **2** and **6-11** all three variables were determined from a global analysis. The 95 %-confidence intervals of K_D and ΔH° were calculated with the 1-dimensional error surface projection. ΔG° and $-T\Delta S^\circ$ (change in entropy) were calculated using the equation $\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = -RT \ln K_A$ with T being the absolute temperature and R the universal gas constant ($8.314 \text{ J mol}^{-1} \text{ K}^{-1}$).

Acknowledgements

We sincerely thank Dr. Helene Kovacs from Bruker BioSpin (Fällanden, Switzerland) for acquisition of NMR assignment experiments. B.F. thanks the German Academic Exchange Service (DAAD) for a stipendship.

Keywords: FimH · glycomimetics · ITC · NMR spectroscopy · urinary tract infections · X-ray diffraction

- [1] B. Foxman, *Am J Med* **2002**, 113 Suppl 1A, 5S-13S.
- [2] T. J. Wiles, R. R. Kulesus, M. A. Mulvey, *Exp Mol Pathol* **2008**, 85, 11-19.
- [3] a) G. C. Schito, K. G. Naber, H. Botto, J. Palou, T. Mazzel, L. Gualco, A. Marchese, *Int J Antimicrob Agents* **2009**, 34, 407-413; b) G. G. Zhanel, T. L. Hisanaga, N. M. Laing, M. R. DeCorby, K. A. Nichol, B. Weshnoweski, J. Johnson, A. Noreddin, D. E. Low, J. A. Karlowsky, N. Group, D. J. Hoban, *Int J Antimicrob Agents* **2006**, 27, 468-475.
- [4] a) E. V. Sokurenko, V. Chesnokova, R. J. Doyle, D. L. Hasty, *J Biol Chem* **1997**, 272, 17880-17886; b) N. Sharon, *FEBS Lett* **1987**, 217, 145-157.
- [5] G. Waksman, S. J. Hultgren, *Nat Rev Microbiol* **2009**, 7, 765-774.
- [6] I. Le Trong, P. Aprikian, B. A. Kidd, M. Forero-Shelton, V. Tchesnokova, P. Rajagopal, V. Rodriguez, G. Interlandi, R. Klevit, V. Vogel, R. E. Stenkamp, E. V. Sokurenko, W. E. Thomas, *Cell* **2010**, 141, 645-655.
- [7] G. Zhou, W. J. Mo, P. Sebbel, G. Min, T. A. Neubert, R. Glockshuber, X. R. Wu, T. T. Sun, X. P. Kong, *J Cell Sci* **2001**, 114, 4095-4103.
- [8] a) I. Ofek, D. L. Hasty, N. Sharon, *FEMS Immunol Med Microbiol* **2003**, 38, 181-191; b) R. J. Pieters, *Med Res Rev* **2007**, 27, 796-816.
- [9] M. Aronson, O. Medalia, L. Schori, D. Mirelman, N. Sharon, I. Ofek, *J Infect Dis* **1979**, 139, 329-332.
- [10] a) N. Firon, S. Ashkenazi, D. Mirelman, I. Ofek, N. Sharon, *Infect Immun* **1987**, 55, 472-476; b) T. K. Lindhorst, C. Kieburg, U. Krallmann-Wenzel, *Glycoconj J* **1998**, 15, 605-613; c) J. Bouckaert, J. Berglund, M. Schembri, E. De Genst, L. Cools, M. Wuhler, C. S. Hung, J. Pinkner, R. Stattegard, A. Zavialov, D. Choudhury, S. Langermann, S. J. Hultgren, L. Wyns, P. Klemm, S. Oscarson, S. D. Knight, H. De Greve, *Mol Microbiol* **2005**, 55, 441-455; d) O. Sperling, A. Fuchs, T. K. Lindhorst, *Org Biomol Chem* **2006**, 4, 3913-3922; e) Z. Han, J. S. Pinkner, B. Ford, R. Obermann, W. Nolan, S. A. Wildman, D. Hobbs, T. Ellenberger, C. K. Cusumano, S. J. Hultgren, J. W. Janetka, *J Med Chem* **2010**, 53, 4779-4792; f) T. Klein, D. Abgottspon, M. Wittwer, S. Rabbani, J. Herold, X. Jiang, S. Kleeb, C. Luthi, M. Scharenberg, J. Bezencon, E. Gubler, L. Pang, M. Smiesko, B. Cutting, O. Schwardt, B. Ernst, *J Med Chem* **2010**, 53, 8627-8641; g) O. Schwardt, S. Rabbani, M. Hartmann, D. Abgottspon, M. Wittwer, S. Kleeb, A. Zalewski, M. Smiesko, B. Cutting, B. Ernst, *Bioorg Med Chem* **2011**, 19, 6454-6473; h) X. Jiang, D. Abgottspon, S. Kleeb, S. Rabbani, M. Scharenberg, M. Wittwer, M. Haug, O. Schwardt, B. Ernst, *J Med Chem* **2012**, 55, 4700-4713; i) L. Pang, S. Kleeb, K. Lemme, S. Rabbani, M. Scharenberg, A. Zalewski, F. Schadler, O. Schwardt, B. Ernst, *ChemMedChem* **2012**, 7, 1404-1422; j) M. Scharenberg, O. Schwardt, S. Rabbani, B. Ernst, *J Med Chem* **2012**, 55, 9810-9816; k) Z. Han, J. S. Pinkner, B. Ford, E. Chorell, J. M. Crowley, C. K. Cusumano, S. Campbell, J. P. Henderson, S. J. Hultgren, J. W. Janetka, *J Med Chem* **2012**, 55, 3945-3959; l) M. Hartmann, T. K. Lindhorst, *Eur J Org Chem* **2011**, 3583-3609; m) S. Kleeb, L. Pang, K. Mayer, D. Eris, A. Sigl, R. C. Preston, P. Zihlmann, D. Abgottspon, A. Hutter, M. Scharenberg, X. Jiang, G. Navarra, S. Rabbani, M. Smiesko, N. Lüdin, R. P. Jakob, O. Schwardt, T. Maier, T. Sharpe, B. Ernst, **2014**, submitted.
- [11] C. S. Hung, J. Bouckaert, D. Hung, J. Pinkner, C. Widberg, A. DeFusco, C. G. Augustine, R. Strouse, S. Langermann, G. Waksman, S. J. Hultgren, *Mol Microbiol* **2002**, 44, 903-915.
- [12] A. Wellens, C. Garofalo, H. Nguyen, N. Van Gerven, R. Stattegard, J. P. Hernalsteens, L. Wyns, S. Oscarson, H. De Greve, S. Hultgren, J. Bouckaert, *PLoS One* **2008**, 3, e2040.
- [13] S. Rabbani, X. Jiang, O. Schwardt, B. Ernst, *Anal Biochem* **2010**, 407, 188-195.
- [14] a) D. Abgottspon, G. Rolli, L. Hosch, A. Steinhuber, X. Jiang, O. Schwardt, B. Cutting, M. Smiesko, U. Jenal, B. Ernst, A. Trampuz, *J Microbiol Methods* **2010**, 82, 249-255; b) M. Scharenberg, D. Abgottspon, E. Cicek, X. Jiang, O. Schwardt, S. Rabbani, B. Ernst, *Assay Drug Dev Technol* **2011**, 9, 455-464.
- [15] D. Abgottspon, B. Ernst, *Chimia* **2012**, 66, 166-169.
- [16] a) A. Clery, M. Schubert, F. H. Allain, *Chimia* **2012**, 66, 741-746; b) M. del Carmen Fernandez-Alonso, D. Diaz, M. A. Berbis, F. Marcelo, J. Canada, J. Jimenez-Barbero, *Curr Protein Pept Sci* **2012**, 13, 816-830.
- [17] a) S. B. Shuker, P. J. Hajduk, R. P. Meadows, S. W. Fesik, *Science* **1996**, 274, 1531-1534; b) L. Skjaerven, L. Codutti, A. Angelini, M. Grimaldi, D. Latek, P. Monecke, M. K. Dreyer, T. Carlomagno, *J Am Chem Soc* **2013**, 135, 5819-5827.
- [18] a) G. Wagner, A. Pardi, K. Wuthrich, *J Am Chem Soc* **1983**, 105, 5948-5949; b) M. P. Williamson, *Prog Nucl Magn Reson Spectrosc* **2013**, 73, 1-16.
- [19] S. Vanwetswinkel, A. N. Volkov, Y. G. Sterckx, A. Garcia-Pino, L. Buts, W. F. Vranken, J. Bouckaert, R. Roy, L. Wyns, N. A. van Nuland, *J Med Chem* **2014**, 57, 1416-1427.
- [20] A. Wellens, M. Lahmann, M. Touaibia, J. Vaucher, S. Oscarson, R. Roy, H. Remaut, J. Bouckaert, *Biochem* **2012**, 51, 4790-4799.
- [21] G. Roos, A. Wellens, M. Touaibia, N. Yamakawa, P. Geerlings, R. Roy, L. Wyns, J. Bouckaert, *ACS Med Chem Lett* **2013**, 4, 1085-1090.
- [22] D. Choudhury, A. Thompson, V. Stojanoff, S. Langermann, J. Pinkner, S. J. Hultgren, S. D. Knight, *Science* **1999**, 285, 1061-1066.
- [23] M. Sattler, J. Schleucher, C. Griesinger, *Prog Nucl Magn Reson Spectrosc* **1999**, 34, 93-158.
- [24] The backbone resonance assignment was deposited in the Biological Magnetic Resonance Database (BMRB) entry: 12345
- [25] M. Scharenberg, X. Jiang, L. Pang, G. Navarra, S. Rabbani, F. Binder, O. Schwardt, B. Ernst, *ChemMedChem* **2014**, 9, 78-83.
- [26] M. J. Cloninger, H. W. Whitlock, *J Org Chem* **1998**, 63, 6153-6159.
- [27] E. A. Meyer, R. K. Castellano, F. Diederich, *Angew Chem Int Ed Engl* **2003**, 42, 1210-1250.
- [28] A. Zalewski, PhD thesis thesis, University of Basel (Basel), **2013**.
- [29] V. B. Rodriguez, B. A. Kidd, G. Interlandi, V. Tchesnokova, E. V. Sokurenko, W. E. Thomas, *J Biol Chem* **2013**, 288, 24128-24139.
- [30] B. W. Sigurskjold, *Anal Biochem* **2000**, 277, 260-266.
- [31] a) L. Frick, C. Yang, V. E. Marquez, R. Wolfenden, *Biochem* **1989**, 28, 9423-9430; b) W. M. Kati, R. Wolfenden, *Biochem* **1989**, 28, 7919-7927.
- [32] R. U. Lemieux, *Acc Chem Res* **1996**, 29, 373-380.
- [33] a) J. M. Myslinski, J. E. DeLorbe, J. H. Clements, S. F. Martin, *J Am Chem Soc* **2011**, 133, 18518-18521; b) K. L. Portman, J. Long, S. Carr, L. Briand, D. J. Winzor, M. S. Searle, D. J. Scott, *Biochem* **2014**, 53, 2371-2379; c) P. W. Snyder, J. Mecinovic, D. T. Moustakas, S. W. Thomas, 3rd, M. Harder, E. T. Mack, M. R. Lockett, A. Heroux, W. Sherman, G. M. Whitesides, *Proc Natl Acad Sci U S A* **2011**, 108, 17889-17894.
- [34] a) J. Lundquist, E. J. Toone, *Chem Rev* **2002**, 102, 555-578; b) W. B. Turnbull, B. L. Precious, S. W. Homans, *J Am Chem Soc* **2004**, 126, 1047-1054.
- [35] M. Ambrosi, N. R. Cameron, B. G. Davis, *Org Biomol Chem* **2005**, 3, 1593-1608.
- [36] W. F. Vranken, W. Boucher, T. J. Stevens, R. H. Fogh, A. Pajon, M. Llinas, E. L. Ulrich, J. L. Markley, J. Ionides, E. D. Laue, *Proteins* **2005**, 59, 687-696.
- [37] M. Pellecchia, P. Sebbel, U. Hermanns, K. Wuthrich, R. Glockshuber, *Nat Struct Biol* **1999**, 6, 336-339.
- [38] a) W. Kabsch, *Acta Crystallogr D Biol Crystallogr* **2010**, 66, 133-144; b) W. Kabsch, *Acta Crystallogr D Biol Crystallogr* **2010**, 66, 125-132.
- [39] A. J. McCoy, *Acta Crystallogr D Biol Crystallogr* **2007**, 63, 32-41.
- [40] P. Emsley, K. Cowtan, *Acta Crystallogr D Biol Crystallogr* **2004**, 60, 2126-2132.
- [41] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *Acta Crystallogr D Biol Crystallogr* **2010**, 66, 213-221.
- [42] D. M. van Aalten, R. Bywater, J. B. Findlay, M. Hendlich, R. W. Hoof, G. Vriend, *J Comput Aided Mol Des* **1996**, 10, 255-262.
- [43] V. B. Chen, W. B. Arendall, 3rd, J. J. Headd, D. A. Keedy, R. M. Immormino, G. J. Kapral, L. W. Murray, J. S. Richardson, D. C. Richardson, *Acta Crystallogr D Biol Crystallogr* **2010**, 66, 12-21.
- [44] J. C. Houtman, P. H. Brown, B. Bowden, H. Yamaguchi, E. Appella, L. E. Samelson, P. Schuck, *Protein Sci* **2007**, 16, 30-42.

Received: ((will be filled in by the editorial staff))

Published online: ((will be filled in by the editorial staff))

Supporting Information

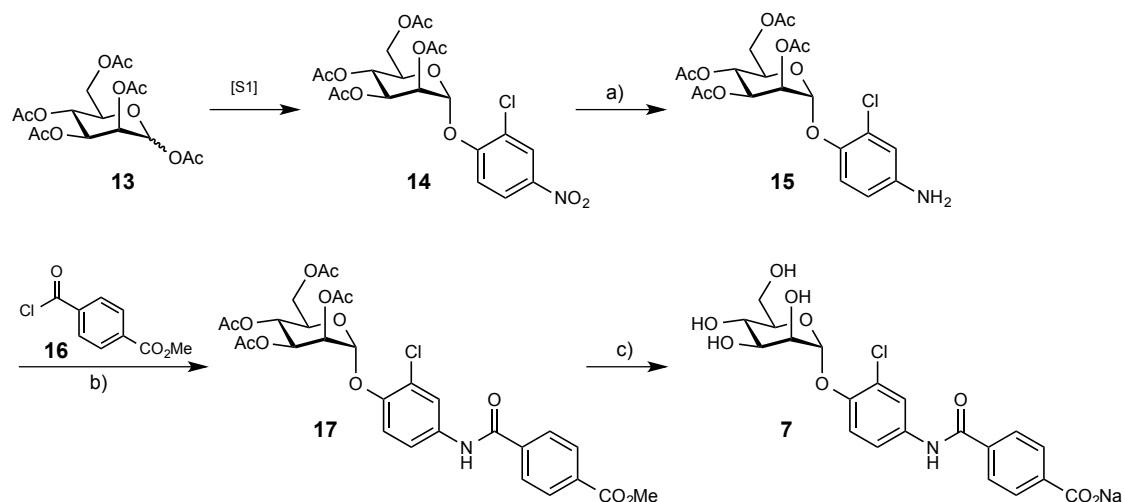
Refined Binding modes of FimH Antagonists – A detailed analysis by NMR and X-ray

Brigitte Fiege,^[a] Roland C. Preston,^[a] Roman P. Jakob,^[b] Pascal Zihlmann,^[a] Said Rabbani,^[a] Oliver Schwardt,^[a] Xiaohua Jiang,^[a] Timm Maier,^[b] and Beat Ernst*^[a]

[a] Institute of Molecular Pharmacy, Pharmacenter, University of Basel, Klingelbergstrasse 50, 4056 Basel (Switzerland)

[b] Structural Biology, Biocenter, University of Basel, Klingelbergstrasse 70, 4056 Basel (Switzerland)

Synthesis of compound 7



Scheme S1. a) H₂ (1 atm), cat. PtO₂, morpholine, MeOH, rt, 2 h (94%); b) cat. DMAP, pyridine, 0 °C to rt, 19 h (96%); c) i. NaOMe/MeOH, rt, 3 h; ii. NaOH, H₂O/dioxane, rt, 16 h (56%).

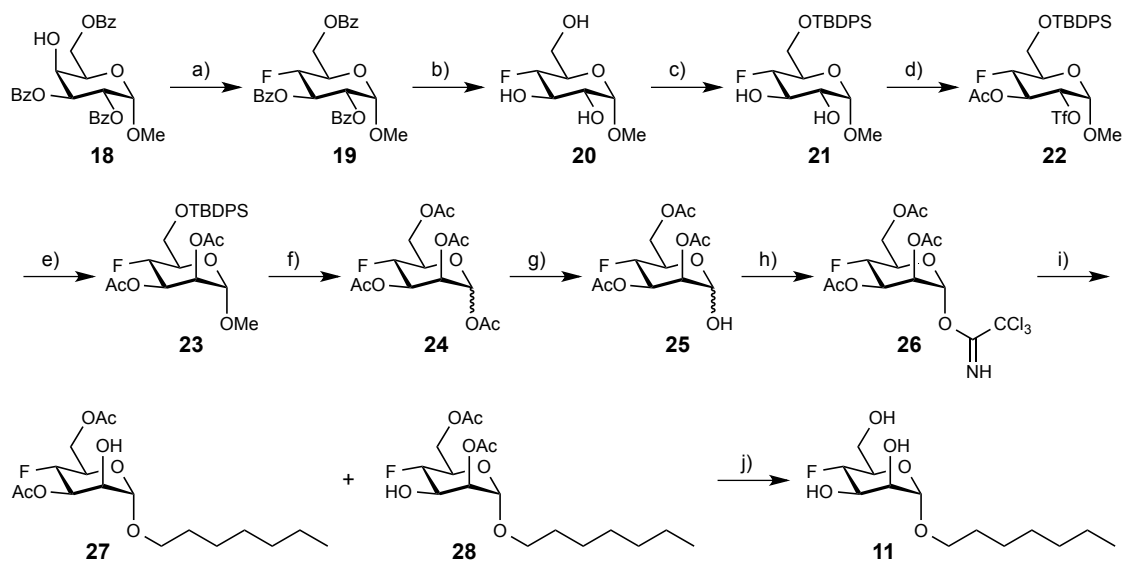
4-Amino-2-chlorophenyl 2,3,4,6-tetra-O-acetyl- α -D-mannopyranoside (15). A suspension of **14**^[S1] (2.00 g, 3.97 mmol), morpholine (200 μ L) and PtO₂ (100 mg) in MeOH (20 mL) was hydrogenated (1 atm H₂) for 2 h. Then the mixture was filtered and concentrated *in vacuo*. The residue was purified by MPLC on silica (petroleum ether/EtOAc) to yield **15** (1.76 g, 94%) as colorless foam. ¹H NMR (500 MHz, CDCl₃): δ = 2.03, 2.06, 2.06, 2.18 (4 s, 12H, 4 COCH₃), 4.10 (dd, J = 1.8, 11.7 Hz, 1H, H-6a), 4.24 (ddd, J = 2.0, 5.4, 9.9 Hz, 1H, H-5), 4.28 (dd, J = 5.4, 11.8 Hz, 1H, H-6b), 5.36 (t, J = 10.0 Hz, 1H, H-4), 5.39 (d, J = 1.7 Hz, 1H, H-1), 5.51 (dd, J = 1.9, 3.4 Hz, 1H, H-2), 5.58 (dd, J = 3.5, 10.0 Hz, 1H, H-3), 6.69 (dd, J = 2.6, 8.7 Hz, 1H, Ar-H), 6.91 (d, J = 2.6 Hz, 1H, Ar-H), 7.01 (d, J = 8.8 Hz, 1H, Ar-H); ¹³C NMR (125 MHz, CDCl₃): δ = 20.7, 20.7, 20.9 (4C, 4 COCH₃), 62.2 (C-6), 65.9 (C-4), 68.8 (C-3), 69.4 (C-2), 69.6 (C-5), 97.6 (C-1), 116.0, 118.7, 119.2, 125.6, 145.8 (6C, Ar-C), 169.8, 169.9, 170.5 (4C, 4 CO); ESI-MS: m/z : Calcd for C₂₀H₂₅ClNO₁₀ [M+H]⁺: 474.1, found: 474.1.

Methyl 4-[(4-(2,3,4,6-tetra-O-acetyl- α -D-mannopyranosyloxy)-3-chlorophenyl)carbamoyl]-benzoate (17). To a solution of **15** (200 mg, 0.422 mmol) and DMAP (10 mg) in pyridine (10 mL) was added methyl 4-chlorocarbonylbenzoate (**16**, 126 mg, 0.636 mmol) at 0 °C under argon. The mixture was stirred at rt for 19 h and then concentrated. The remains were dissolved in DCM (20 mL) and subsequently washed with 0.1 N aq. HCl (10 mL) and satd aq. NaHCO₃ (10 mL). The organic layer was dried on Na₂SO₄, filtered and concentrated. The residue was purified by MPLC on silica (petroleum ether/EtOAc) to afford **17** (258 mg, 96%) as a colorless foam. $[\alpha]_D^{+25}$ +63.0 (c 1.38, CHCl₃); ¹H NMR (500 MHz, CDCl₃): δ = 2.02, 2.05, 2.06, 2.19 (4 s, 12H, 4 COCH₃), 3.94 (s, 3H, OMe), 4.09 (dd, J = 2.0, 12.1 Hz, 1H, H-6a), 4.20 (ddd, J = 2.0, 5.2, 9.9 Hz, 1H, H-5), 4.28 (dd, J = 5.3, 12.1 Hz, 1H, H-6b), 5.38 (t, J = 10.1 Hz, 1H, H-4), 5.50 (s, 1H, H-1), 5.52 (dd, J = 1.7, 3.3 Hz, 1H, H-2), 5.58 (dd, J = 3.4, 10.0 Hz, 1H, H-3), 7.15 (d, J = 8.9 Hz, 1H, C₆H₃), 7.53 (dd, J = 1.9, 8.9 Hz, 1H, C₆H₃), 7.77 (d, J = 2.3 Hz, 1H, C₆H₃), 7.90 (d, J = 8.3 Hz, 2H, C₆H₄), 8.10 (d, J = 8.0 Hz, 2H, C₆H₄), 8.15 (s,

1H, NH); ¹³C NMR (125 MHz, CDCl₃): δ = 20.6, 20.7, 20.8, 21.0 (4 COCH₃), 52.5 (OMe), 62.1 (C-6), 65.8 (C-4), 68.8 (C-3), 69.3 (C-2), 69.7 (C-5), 97.0 (C-1), 117.7, 119.8, 122.6, 124.8, 127.1, 130.0, 133.1, 133.8, 138.3, 148.2 (12C, C₆H₃, C₆H₄), 164.8 (CONH), 166.1 (CO₂Me), 169.7, 169.8, 170.0, 170.6 (4 COCH₃); ESI-MS: *m/z*: Calcd for C₂₉H₃₀CINNaO₁₃ [M+Na]⁺: 658.1, found: 658.1.

Sodium 4-[(3-chloro-4-(α-D-mannopyranosyloxy)phenyl)carbamoyl]benzoate (7). To a solution of **17** (29.0 mg, 45.6 μmol) in MeOH (4 mL) was added 1 M NaOMe/MeOH (0.4 mL) under argon. The mixture was stirred at rt for 3 h and then concentrated *in vacuo*. The residue was dissolved in H₂O/dioxane (4 mL, 1:1) and treated with 1 M aq. NaOH (0.4 mL) for 16 h. The mixture was neutralized with 1 M aq. HCl and evaporated. The residue was purified by preparative LC-MS (RP-18, H₂O/MeCN + 0.2% HCO₂H) to give the acid form of **17**. The acid was dissolved in 0.1 M aq. NaOH (1 mL) and re-purified by size-exclusion chromatography (P2 column, H₂O) to yield **17** (11.5 mg, 56%) as a white powder after a final lyophilization from H₂O. [α]_D +63.8 (c 0.27, MeOH); ¹H NMR (500 MHz, CD₃OD): δ = 3.68 (ddd, *J* = 2.1, 5.4, 9.7 Hz, 1H, H-5), 3.73 (dd, *J* = 5.4, 11.8 Hz, 1H, H-6a), 3.76 (t, *J* = 9.6 Hz, 1H, H-4), 3.80 (dd, *J* = 2.1, 11.8 Hz, 1H, H-6b), 3.98 (dd, *J* = 3.4, 9.4 Hz, 1H, H-3), 4.11 (dd, *J* = 1.8, 3.3 Hz, 1H, H-2), 5.51 (d, *J* = 1.5 Hz, 1H, H-1), 7.37 (d, *J* = 9.0 Hz, 1H, C₆H₃), 7.56 (dd, *J* = 2.5, 8.9 Hz, 1H, C₆H₃), 7.89 (d, *J* = 2.5 Hz, 1H, C₆H₃), 7.99 (d, *J* = 8.4 Hz, 2H, C₆H₄), 8.14 (d, *J* = 8.4 Hz, 2H, C₆H₄); ¹³C NMR (125 MHz, CD₃OD): δ = 62.7 (C-6), 68.3 (C-4), 71.9 (C-2), 72.4 (C-3), 75.9 (C-5), 101.2 (C-1), 118.8, 121.7, 123.9, 124.9, 128.7, 130.9, 135.0, 135.1, 140.0, 150.2 (12C, C₆H₃, C₆H₄), 167.8, 168.9 (CONH, CO₂); HR-MS: *m/z*: Calcd for C₂₀H₂₀CINNaO₉ [M+H]⁺: 476.0719, found: 476.0725.

Synthesis of compound 11



Scheme S2. a) DAST, pyridine, DCM, -30 °C to rt, overnight (66%); b) NaOMe/MeOH, rt, 3 h (97%); c) TBDPSCI, imidazole, DMF, 3.5 h (83%); d) Tf₂O, pyridine, DCM, -18 to -10 °C, 1.5 h, then AcCl, -8 °C to rt, 1 h (96%); e) CsOAc, 18-crown-6, toluene, reflux, 2.5 h (59%); f) 4% H₂SO₄/Ac₂O, rt, overnight (98%); g) Me₂NH/THF, rt, 3 h (70%); h) trichloroacetonitrile, NaH, DCM (91%); i) *n*-heptanol, TMSOTf, DCM, rt, 2 h (**27**: 25%, **28**: 32%); j) NaOMe/MeOH, rt, overnight (92%).

Methyl 2,3,6-tri-O-benzoyl-4-deoxy-4-fluoro-α-D-glucopyranoside (19).^[S2] A solution of methyl 2,3,6-tri-O-benzoyl-α-D-galactopyranoside (**18**,^[S2] 5.06 g, 10.0 mmol) in dry DCM (10.0 mL) was added to a mixture of pyridine (2.5 mL) and DAST (1.7 mL) in dry DCM (25 mL) at -30 °C. After stirring at rt overnight, the reaction mixture was cooled to 0 °C, quenched with MeOH and concentrated. The residue was purified by MPLC on silica (petroleum ether/EtOAc, 15:1-9:1) to afford **19** (3.36 g, 66%) as white foam. ¹H NMR (500 MHz, CDCl₃): δ = 3.47 (s, 3H, OCH₃), 4.33 (m, 1H, H-5), 4.63 (ddd, *J* = 1.0, 5.0, 12.0 Hz, 1H, H-6a), 4.72 (m, 1H, H-6b), 4.77 (dt, *J* = 9.5, 51.0 Hz, 1H, H-4), 5.18 (t, *J* = 3.5 Hz, 1H, H-1), 5.21 (dd, *J* = 3.5, 10.0 Hz, 1H, H-2), 6.14 (ddd, *J* = 9.5, 10.0, 14.5 Hz, 1H, H-3), 7.38 (m, 4H, Ar-H), 7.50 (m, 4H, Ar-H), 7.61 (m, 1H, Ar-H), 8.00 (m, 4H, Ar-H), 8.10 (m, 2H, Ar-H); ¹⁹F NMR (470 MHz, CDCl₃): δ = -197.0.

Methyl 4-deoxy-4-fluoro-α-D-glucopyranoside (20). To a suspension of **19** (3.36 g, 6.60 mmol) in dry MeOH (30 mL) was added 0.5 M NaOMe/MeOH (1.98 mL, 0.99 mmol) at rt. The reaction mixture was stirred at rt for 3 h, then neutralized with Amberlyst-15, filtered and concentrated. The residue was purified by MPLC on silica (DCM/MeOH, 20:1-9:1) to afford **20** (1.26 g, 97%) as a white solid, which was directly used in the next step.

Methyl 6-O-*tert*-butyldiphenylsilyl-4-deoxy-4-fluoro-α-D-glucopyranoside (21). To a solution of **20** (227 mg, 1.16 mmol) and imidazole (177 mg, 2.6 mmol) in dry DMF was added TBDPSCI (0.34

mL, 1.30 mmol) at rt. The reaction mixture was stirred for 3.5 h at rt under argon, then diluted with DCM, washed with 0.1 N HCl and brine, dried over Na₂SO₄, filtered and concentrated *in vacuo*. The residue was purified by MPLC on silica (DCM/MeOH, 15:1-9:1) to provide **21** (422 mg, 83%) as a colorless oil. $[\alpha]_D^{+67.9}$ (c 1.80, CHCl₃); ¹H NMR (500 MHz, CDCl₃): δ = 1.05 (s, 9H, C(CH₃)₃), 2.33 (d, J = 9.0 Hz, 1H, 2-OH), 2.86 (m, 1H, 3-OH), 3.57 (td, J = 4.0, 9.5 Hz, 1H, H-2), 3.46 (s, 3H, OCH₃), 3.77 (m, 1H, H-5), 3.88 (m, 2H, H-6), 3.96 (m, 1H, H-3), 4.43 (dt, J = 9.5, 51.0 Hz, 1H, H-4), 4.80 (t, J = 3.5 Hz, 1H, H-1), 7.37-7.44 (m, 6H, Ar-H), 7.69 (m, 4H, Ar-H); ¹³C NMR (125 MHz, CDCl₃): δ = 19.3 (C(CH₃)₃), 26.7 (C(CH₃)₃), 55.3 (OCH₃), 62.3 (C-6), 69.8 (d, J = 23.1 Hz, C-5), 72.2 (d, J = 7.8 Hz, C-2), 73.2 (d, J = 17.4 Hz, C-3), 88.8 (d, J = 180.8 Hz, C-4), 98.7 (d, J = 1.3 Hz, C-1), 127.6, 127.7, 129.7, 133.2, 133.2, 135.6, 135.6 (12C, Ar-C); ¹⁹F NMR (470 MHz, CDCl₃): δ = -198.9; ESI-MS: m/z : Calcd for C₂₃H₃₁FNaO₅Si [M+Na]⁺: 457.2, found: 457.1.

Methyl 3-O-acetyl-6-O-tert-butylidiphenylsilyl-4-deoxy-4-fluoro-2-O-trifluoromethylsulfonyl- α -D-glucopyranoside (22). To a solution of **21** (1.65 g, 3.79 mmol) and pyridine (1.23 mL, 15.2 mmol) in dry DCM (35 mL) was added trifluoromethanesulfonic anhydride (0.74 mL, 4.36 mmol) slowly at -18 °C. The reaction mixture was stirred at -15 to -10 °C for 1.5 h, then acetyl chloride (0.323 mL) was added at -8 °C. The cooling bath was removed and the reaction mixture was stirred for 1 h at rt, then diluted with DCM, and washed with 5% NaHCO₃ and brine. The organic layer was dried over Na₂SO₄ and concentrated *in vacuo*. The residue was purified by MPLC on silica (petroleum ether/EtOAc, 15:1-9:1) to afford **22** (2.23 g, 96%) as a syrup. $[\alpha]_D^{+64.5}$ (c 1.58, CHCl₃); ¹H NMR (500 MHz, CDCl₃): δ = 1.07 (s, 9H, C(CH₃)₃), 2.16 (s, 3H, COCH₃), 3.47 (s, 3H, OCH₃), 3.90 (m, 3H, H-5, H-6), 4.58 (dt, J = 9.5, 50.5 Hz, 1H, H-4), 4.67 (dd, J = 3.5, 10.0 Hz, 1H, H-2), 4.99 (t, J = 3.5 Hz, 1H, H-1), 5.72 (dt, J = 9.5, 13.5 Hz, 1H, H-3), 7.42 (m, 6H, Ar-H), 7.68 (m, 4H, Ar-H); ¹³C NMR (125 MHz, CDCl₃): δ = 20.5 (COCH₃), 19.3 (C(CH₃)₃), 26.7 (C(CH₃)₃), 55.8 (OCH₃), 61.7 (C-6), 68.9 (d, J = 21.3 Hz, C-3), 69.4 (d, J = 23.1 Hz, C-5), 80.9 (d, J = 7.9 Hz, C-2), 86.5 (d, J = 187.6 Hz, C-4), 96.5 (C-1), 127.7, 127.8, 129.8, 132.9, 132.9, 135.6, 135.6 (12C, Ar-C), 169.4 (CO); ¹⁹F NMR (470 MHz, CDCl₃): δ = -197.9, -74.7; ESI-MS: m/z : Calcd for C₂₆H₃₂F₄NaO₈Si [M+Na]⁺: 631.1, found: 631.1.

Methyl 2,3-di-O-acetyl-6-O-tert-butylidiphenylsilyl-4-deoxy-4-fluoro- α -D-mannopyranoside (23). A mixture of **22** (2.07 g, 3.39 mmol), CsOAc (977 mg, 5.08 mmol) and 18-crown-6 (1.34 g, 5.08 mmol) in dry toluene (20 mL) was refluxed for 2.5 h. After cooling to rt, toluene was removed *in vacuo* and the residue was dissolved in EtOAc, washed with water and brine, and dried over Na₂SO₄. The solvent was removed *in vacuo* and the residue was purified by MPLC on silica (petroleum ether/EtOAc, 9:1-6:1) to afford **23** (1.03 g, 59%) as a white solid. $[\alpha]_D^{+27.3}$ (c 0.63, CHCl₃); ¹H NMR (500 MHz, CDCl₃): δ = 1.10 (s, 9H, C(CH₃)₃), 2.09, 2.13 (2 s, 6H, 2 COCH₃), 3.40 (s, 3H, OCH₃), 3.84 (m, 1H, H-5), 3.92 (m, 1H, H-6a), 3.99 (m, 1H, H-6b), 4.73 (s, 1H, H-1), 4.96 (dt, J = 9.5, 51.5 Hz, 1H, H-4), 5.29 (d, J = 1.5 Hz, 1H, H-2), 5.45 (ddd, J = 3.5, 9.5, 13.5 Hz, 1H, H-3), 7.41 (m, 6H, Ar-H), 7.71-7.75 (m, 4H, Ar-H); ¹³C NMR (125 MHz, CDCl₃): δ = 20.74, 20.82 (2 COCH₃), 19.4 (C(CH₃)₃), 26.7 (C(CH₃)₃), 55.1 (OCH₃), 62.0 (C-6), 69.4 (d, J = 18.9 Hz, C-3), 70.4 (d, J = 8.4 Hz, C-2), 70.6 (d, J = 23.6 Hz, C-5), 84.8 (d, J = 179.9 Hz, C-4), 98.25 (C-1), 127.6, 127.7, 129.7, 129.7, 133.0, 133.4,

135.5, 135.7 (12C, Ar-C), 169.9, 170.0 (2 CO); ^{19}F NMR (470 MHz, CDCl_3): δ = -204.6; HR-MS: m/z : Calcd for $\text{C}_{27}\text{H}_{35}\text{FNaO}_7\text{Si}$ $[\text{M}+\text{Na}]^+$: 541.2034, found: 541.2028.

1,2,3,6-Tetra-O-acetyl-4-deoxy-4-fluoro- α -D-mannopyranoside (24). To a solution of **23** (940 mg, 1.54 mmol) in acetic anhydride (8.0 mL) was added H_2SO_4 (95-97%, 0.35 mL) at 0 °C. The reaction mixture was stirred at rt overnight, then diluted with ethyl acetate and neutralized carefully with satd. aq. NaHCO_3 at 0 °C. The aqueous layer was extracted twice with EtOAc, the combined organic layers were washed with brine and dried over Na_2SO_4 . The solvent was removed *in vacuo* and the residue was purified by MPLC on silica (petroleum ether/EtOAc, 3:1-7:3) to afford **24** (530 mg, 98%) as a syrup. $[\alpha]_{\text{D}} +27.2$ (c 0.58, CHCl_3); ^1H NMR (500 MHz, CDCl_3): δ = 2.07, 2.10, 2.14, 2.17 (4 s, 12H, 4 COCH_3), 4.10 (m, 1H, H-5), 4.29 (ddd, J = 1.0, 5.0, 12.0 Hz, 1H, H-6a), 4.38 (dt, J = 2.0, 12.5 Hz, 1H, H-6b), 4.74 (dt, J = 10.0, 51.5 Hz, 1H, H-4), 5.28 (dd, J = 3.0, 5.0 Hz, 1H, H-2), 5.42 (ddd, J = 3.5, 9.5, 13.0 Hz, 1H, H-3), 6.04 (t, J = 3.0 Hz, 1H, H-1); ^{13}C NMR (125 MHz, CDCl_3): δ = 20.6, 20.6, 20.7, 20.8 (4 COCH_3), 62.1 (C-6), 68.7 (d, J = 18.9 Hz, C-3), 69.9 (d, J = 8.3 Hz, C-2), 70.0 (d, J = 23.8 Hz, C-5), 84.7 (d, J = 182.4 Hz, C-4), 90.3 (d, J = 0.5 Hz, C-1), 168.0, 169.5, 169.8, 170.5 (4 C=O); ^{19}F NMR (470 MHz, CDCl_3): δ = -204.5; HR-MS: m/z : Calcd for $\text{C}_{14}\text{H}_{19}\text{FNaO}_9$ $[\text{M}+\text{Na}]^+$: 373.0911, found: 373.0904.

2,3,6-Tri-O-acetyl-4-deoxy-4-fluoro- α -D-mannopyranoside (25). To a solution of **24** (530 mg, 1.51 mmol) in dry THF (12.0 mL) was added Me_2NH (1.51 mL, 2 M in THF, 0.89 mmol). The reaction mixture was stirred for 3 h at rt and then concentrated. The residue was purified by MPLC on silica (petroleum ether/EtOAc, 3:1-3:2) to afford **25** (326 mg, 70%) as a pale yellow solid. $[\alpha]_{\text{D}} -7.9$ (c 0.77, CHCl_3); ^1H NMR (500 MHz, CDCl_3): δ = 2.08, 2.12, 2.14 (3 s, 9H, 3 COCH_3), 3.49 (s, 1H, 1-OH), 4.29 (m, 2H, H-5, H-6a), 4.44 (m, 1H, H-6b), 4.71 (dt, J = 10.0, 52.0 Hz, 1H, H-4), 5.21 (s, 1H, H-1), 5.30 (d, J = 1.5 Hz, 1H, H-2), 5.50 (ddd, J = 3.5, 10.0, 13.5 Hz, 1H, H-3); ^{13}C NMR (125 MHz, CDCl_3): δ = 20.7, 20.8, 20.8 (3 COCH_3), 62.5 (C-6), 68.0 (d, J = 23.5 Hz, C-5), 68.7 (d, J = 18.9 Hz, C-3), 70.6 (d, J = 8.3 Hz, C-2), 85.2 (d, J = 181.9 Hz, C-4), 92.1 (C-1), 169.9, 169.9, 170.8 (3 CO); ^{19}F NMR (470 MHz, CDCl_3): δ = -204.3; ESI-MS: m/z : Calcd for $\text{C}_{12}\text{H}_{17}\text{FNaO}_8$ $[\text{M}+\text{Na}]^+$: 331.1, found: 331.0.

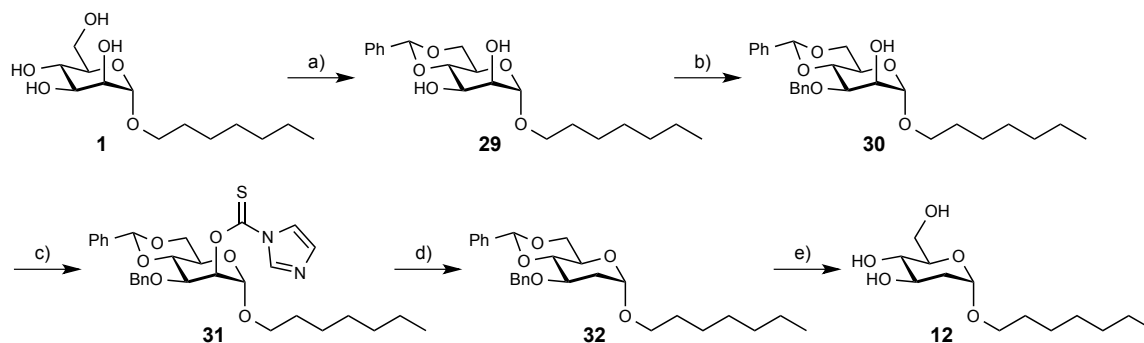
2,3,6-Tri-O-acetyl-4-deoxy-4-fluoro- α -D-mannopyranosyl trichloroacetimidate (26). To a solution of **25** (187 mg, 0.607 mmol) and trichloroacetonitrile (0.61 mL, 6.07 mmol) in dry DCM (5 mL) was added NaH (60%, 3.6 mg) at rt. The reaction mixture was stirred at rt for 1 h, then silica gel was added and the solvent was removed *in vacuo*. The residue was purified by MPLC on silica (petroleum ether/EtOAc, 9:1-3:1) to afford **26** (250 mg, 91%) as colorless syrup. $[\alpha]_{\text{D}} +40.0$ (c 0.29, CHCl_3); ^1H NMR (500 MHz, CDCl_3): δ = 2.09, 2.10, 2.18 (3 s, 9H, 3 COCH_3), 4.26 (m, 1H, H-5), 4.31 (ddd, J = 1.5, 5.0, 12.0 Hz, 1H, H-6a), 4.46 (dt, J = 1.5, 12.0 Hz, 1H, H-6b), 4.81 (dt, J = 9.5, 51.5 Hz, 1H, H-4), 5.50 (m, 2H, H-2, H-3), 6.25 (s, 1H, H-1), 8.80 (s, 1H, NH); ^{13}C NMR (125 MHz, CDCl_3): δ = 20.6, 20.7 (3C, 3 COCH_3), 62.0 (C-6), 68.4 (d, J = 8.4 Hz, C-2), 68.7 (d, J = 18.9 Hz, C-3), 70.6 (d, J = 23.9 Hz,

C-5), 84.5 (d, $J = 182.8$ Hz, C-4), 94.3 (C-1), 159.7 (OCN), 169.5, 169.6, 170.5 (3 CO); ^{19}F NMR (470 MHz, CDCl_3): $\delta = -204.5$; ESI-MS: m/z : Calcd for $\text{C}_{14}\text{H}_{17}\text{Cl}_3\text{FNNaO}_8$ $[\text{M}+\text{Na}]^+$: 474.0, found: 473.9.

***n*-Heptyl 3,6-di-*O*-acetyl-4-deoxy-4-fluoro- α -D-mannopyranoside (27) and *n*-heptyl 2,6-di-*O*-acetyl-4-deoxy-4-fluoro- α -D-mannopyranoside (28).** To a mixture of **26** (116 mg, 0.265 mmol) and *n*-heptanol (75 μL , 0.53 mmol) in dry DCM (3 mL) was added TMSOTf (6.0 μL , 0.032 mmol) at rt. The reaction mixture was stirred at rt for 2 h, then neutralized with Et_3N and concentrated to dryness. The residue was purified by MPLC on silica (petroleum ether/EtOAc, 6:1-2:1) to afford **27** (24 mg, 25%) and **28** (31 mg, 32%) as colorless oils. **27**: $[\alpha]_{\text{D}} +39.8$ (c 1.13, CHCl_3); ^1H NMR (500 MHz, CDCl_3): $\delta = 0.88$ (t, $J = 7.0$ Hz, 3H, H-7'), 1.28 (m, 8H, H-3', H-4', H-5', H-6'), 1.59 (m, 2H, H-2'), 2.10, 2.16 (2 s, 6H, 2 COCH_3), 2.27 (d, $J = 3.5$ Hz, 1H, 2-OH), 3.44 (m, 1H, H-1'a), 3.68 (m, 1H, H-1'b), 4.02 (m, 1H, H-5), 4.08 (s, 1H, H-2), 4.29 (dd, $J = 4.5, 12.0$ Hz, 1H, H-6a), 4.38 (d, $J = 12.0$ Hz, 1H, H-6b), 4.72 (dt, $J = 10.0, 52.0$ Hz, 1H, H-4), 4.81 (s, 1H, H-1), 5.33 (t, $J = 12.0$ Hz, 1H, H-3); ^{13}C NMR (125 MHz, CDCl_3): $\delta = 14.0$ (C-7'), 20.8, 21.0 (2 COCH_3), 22.6, 26.0, 29.0, 29.2, 31.7 (C-2', C-3', C-4', C-5', C-6'), 62.6 (C-6), 67.8 (d, $J = 23.4$ Hz, C-5), 68.4 (C-1'), 70.0 (d, $J = 8.0$ Hz, C-2), 71.6 (d, $J = 17.8$ Hz, C-3), 85.4 (d, $J = 181.5$ Hz, C-4), 99.4 (C-1), 169.8, 170.8 (2 CO); ^{19}F NMR (470 MHz, CDCl_3): $\delta = -205.7$; ESI-MS: m/z : Calcd for $\text{C}_{17}\text{H}_{29}\text{FNaO}_7$ $[\text{M}+\text{Na}]^+$: 387.2, found: 387.1. **28**: $[\alpha]_{\text{D}} +35.9$ (c 1.28, CHCl_3); ^1H NMR (500 MHz, CDCl_3): $\delta = 0.88$ (t, $J = 7.0$ Hz, 3H, H-7'), 1.28 (m, 8H, H-3', H-4', H-5', H-6'), 1.56 (m, 2H, H-2'), 2.10, 2.13 (2 s, 6H, 2 COCH_3), 2.53 (d, $J = 4.0$ Hz, 1H, 3-OH), 3.43 (m, 1H, H-1'a), 3.65 (m, 1H, H-1'b), 3.94 (m, 1H, H-5), 4.23 (m, 1H, H-3), 4.28 (dd, $J = 5.5, 12.0$ Hz, 1H, H-6a), 4.38 (d, $J = 12.0$ Hz, 1H, H-6b), 4.54 (dt, $J = 9.5, 52.0$ Hz, 1H, H-4), 4.79 (s, 1H, H-1), 5.14 (s, 1H, H-2); ^{13}C NMR (125 MHz, CDCl_3): $\delta = 14.0$ (C-7'), 20.8, 20.9 (2 COCH_3), 22.6, 26.0, 29.0, 29.2, 31.7 (C-2', C-3', C-4', C-5', C-6'), 62.7 (C-6), 67.6 (d, $J = 23.4$ Hz, C-5), 68.2 (d, $J = 19.0$ Hz, C-3), 68.4 (C-1'), 72.6 (d, $J = 9.0$ Hz, C-2), 88.5 (d, $J = 178.3$ Hz, C-4), 97.3 (C-1), 170.6, 170.7 (2 CO); ^{19}F NMR (470 MHz, CDCl_3): $\delta = -204.5$; HR-MS: m/z : calcd for $\text{C}_{17}\text{H}_{29}\text{FNaO}_7$ $[\text{M}+\text{Na}]^+$: 387.1795, found: 387.1796.

***n*-Heptyl 4-deoxy-4-fluoro- α -D-mannopyranoside (11).** To a mixture of **27** and **28** (118 mg, 0.324 mmol) in dry MeOH (5 mL) was added 0.5 M NaOMe/MeOH (97 μL) at rt. The reaction mixture was stirred overnight, neutralized with Amberlyst-15, filtered and concentrated to dryness. The residue was purified by MPLC on silica (DCM/MeOH, 20:1) to afford **11** (83.4 mg, 92%) as a white solid. $[\alpha]_{\text{D}} +61.5$ (c 1.49, CHCl_3); ^1H NMR (500 MHz, CD_3OD): $\delta = 0.91$ (t, $J = 7.0$ Hz, 3H, H-7'), 1.34 (m, 8H, H-3', H-4', H-5', H-6'), 1.60 (m, 2H, H-2'), 3.43 (m, 1H, H-1'a), 3.66-3.73 (m, 4H, H-5, H-6, H-1'b), 3.83 (m, 1H, H-2), 3.91 (m, 1H, H-3), 4.51 (dt, $J = 9.5, 52.5$ Hz, 1H, H-4), 4.74 (s, 1H, H-1); ^{13}C NMR (125 MHz, CD_3OD): $\delta = 14.4$ (C-7'), 23.7, 27.3, 30.2, 30.6, 33.0 (C-2', C-3', C-4', C-5', C-6'), 62.1 (C-6), 68.8 (C-1'), 70.8 (d, $J = 17.5$ Hz, C-3), 72.2 (d, $J = 24.1$ Hz, C-5), 72.9 (d, $J = 9.1$ Hz, C-2), 89.8 (d, $J = 175.3$ Hz, C-4), 101.5 (C-1); ^{19}F NMR (470 MHz, CDCl_3): $\delta = -206.5$; HR-MS: m/z : Calcd for $\text{C}_{13}\text{H}_{25}\text{FNaO}_8$ $[\text{M}+\text{Na}]^+$: 303.1584, found: 303.1579.

Synthesis of compound 12



Scheme S3. a) PhCH(OMe)_2 , p -TsOH, DMF, 50 °C, 5 h (48%); b) Bu_2SnO , toluene, TBAB, BnBr, reflux (91%); c) TCDI, DCE, reflux, overnight (91%); d) Bu_3SnH , toluene, reflux, 6 h (58%); e) i. HOAc/ H_2O (4:1), 80 °C, 1 h; ii. H_2 (4 bar), $\text{Pd(OH)}_2/\text{C}$, rt, 5 h (46%).

***n*-Heptyl 4,6-O-benzylidene- α -D-mannopyranoside (29).** To a solution of **1** (527 mg, 1.89 mmol) in dry DMF (6.0 mL) were added benzylaldehyde dimethylacetate (0.56 mL) and p -TsOH (18 mg) at rt. The reaction mixture was stirred at 50 °C for 5 h, then diluted with DCM and washed with 5% aq. NaHCO_3 and brine. The organic layer was dried over Na_2SO_4 and concentrated. The residue was purified by MPLC on silica (petroleum ether/EtOAc, 3:1-3:2) to yield **29** (650 mg, 94%) as a glassy solid. ^1H NMR (500 MHz, CD_3OD): δ = 0.91 (t, J = 7.0 Hz, 3H, H-7'), 1.36 (m, 8H, H-3', H-4', H-5', H-6'), 1.62 (m, 2H, H-2'), 3.46 (dt, J = 6.5, 9.5 Hz, 1H, H-1'a), 3.41-3.76 (m, 2H, H-5, H-1'b), 3.81 (t, J = 10.0 Hz, 1H, H-6a), 3.88 (m, 1H, H-2), 3.90-3.95 (m, 2H, H-3, H-4), 4.18 (dd, J = 4.5, 10.0 Hz, 1H, H-6b), 4.77 (d, J = 1.0 Hz, 1H, H-1), 5.60 (s, 1H, CHPh), 7.33-7.37 (m, 3H, Ar-H), 7.49-7.51 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, CD_3OD): δ = 14.5 (C-7'), 23.7, 27.3, 30.2, 30.6, 33.0 (C-2', C-3', C-4', C-5', C-6'), 65.3 (C-5), 68.8 (C-1'), 69.6 (C-4), 69.9 (C-6), 72.7 (C-2), 80.2 (C-3), 102.6 (C-1), 103.4 (CHPh), 127.5, 129.0, 129.9, 139.3 (6C, Ar-C); ESI-MS: m/z : Calcd for $\text{C}_{20}\text{H}_{30}\text{NaO}_6$ $[\text{M}+\text{Na}]^+$: 389.2, found: 389.1.

***n*-Heptyl 3-O-benzyl-4,6-O-benzylidene- α -D-mannopyranoside (30).** A mixture of **29** (152 mg, 0.414 mmol) and dibutyltin oxide (112 mg, 0.456 mmol) in dry toluene (8 mL) was refluxed for 6 h and then concentrated to dryness under reduced pressure. The residue was dissolved in dry toluene (8 mL) and tetrabutylammonium bromide (TBAB, 147 mg, 0.456 mmol) and benzyl bromide (59 μL , 0.50 mmol) were added. The mixture was stirred at 95 °C overnight, concentrated to dryness, and the residue was purified by MPLC on silica (petroleum ether/EtOAc, 9:1-4:1) to give **30** (172 mg, 91%) as colorless oil. $[\alpha]_D^{+34.2}$ (c 1.43, MeOH); ^1H NMR (500 MHz, CDCl_3): δ = 0.89 (t, J = 7.0 Hz, 3H, H-7'), 1.29 (m, 8H, H-3', H-4', H-5', H-6'), 1.57 (m, 2H, H-2'), 2.64 (s, 1H, OH), 3.41 (m, 1H, H-1'a), 3.67 (m, 1H, H-1'b), 3.82-3.88 (m, 2H, H-5, H-6a), 3.94 (dd, J = 3.0, 9.5 Hz, 1H, H-3), 4.06-4.14 (m, 2H, H-2, H-4), 4.27 (m, 1H, H-6b), 4.73 (d, J = 11.5 Hz, 1H, CH_2Ph), 4.86 (m, 1H, H-1), 4.87 (d, J = 11.5 Hz, 1H, CH_2Ph), 5.62 (s, 1H, CHPh), 7.33-7.40 (m, 8H, Ar-H), 7.49-7.51 (m, 2H, Ar-H); ^{13}C NMR (125 MHz, CDCl_3): δ = 14.1 (C-7'), 22.6, 26.0, 29.1, 29.4, 32.8 (C-2', C-3', C-4', C-5', C-6'), 63.2 (C-5), 68.0

(C-1'), 68.9 (C-6), 70.1 (C-2), 73.0 (OCH₂Ph), 75.8 (C-3), 78.9 (C-4), 99.9 (C-1), 101.5 (CHPh), 126.0, 127.8, 127.9, 128.2, 128.5, 128.9, 137.5, 138.1 (12C, Ar-C); HR-MS: *m/z*: Calcd for C₂₇H₃₆NaO₆ [M+Na]⁺: 479.2410, found: 479.2414.

***n*-Heptyl 3-O-benzyl-4,6-O-benzylidene-2-O-(1*H*-imidazole-1-carbonothioyl)- α -D-mannopyranoside (31).** A mixture of **30** (210 mg, 0.460 mmol) and *N,N'*-thiocarbonyldiimidazole (246 mg, 1.38 mmol) in 1,2-dichloroethane (5 mL) was refluxed overnight, then the solution was concentrated *in vacuo*. The residue was dissolved in DCM and washed with 1N HCl and brine. The organic layer was dried over Na₂SO₄ and the solvent was removed *in vacuo*. The residue was purified by MPLC on silica (petroleum ether/EtOAc, 3:1-3:2) to afford **31** (238 mg, 91%) as a yellow oil. [α]_D -9.8 (c 0.44, CH₂Cl₂); ¹H NMR (500 MHz, CDCl₃): δ = 0.90 (t, *J* = 7.0 Hz, 3H, H-7'), 1.31 (m, 8H, H-3', H-4', H-5', H-6'), 1.61 (m, 2H, H-2'), 3.46 (dt, *J* = 6.5, 9.5 Hz, 1H, H-1'a), 3.69 (dt, *J* = 6.5, 9.5 Hz, 1H, H-1'b), 3.85 (t, *J* = 10.0 Hz, 1H, H-6a), 3.94 (td, *J* = 4.5, 10.0 Hz, 1H, H-5), 4.04 (t, *J* = 10.0 Hz, 1H, H-4), 4.19 (dd, *J* = 3.5, 10.0 Hz, 1H, H-3), 4.31 (dd, *J* = 4.5, 10.0 Hz, 1H, H-6b), 4.73 (dd, *J* = 12.0 Hz, 2H, CH₂Ph), 5.01 (d, *J* = 1.5 Hz, 1H, H-1), 5.67 (s, 1H, CHPh), 5.90 (dd, *J* = 1.5, 3.5 Hz, 1H, H-2), 7.07 (d, *J* = 0.5 Hz, 1H, Ar-H), 7.27-7.29 (m, 5H, Ar-H), 7.36-7.42 (m, 3H, Ar-H), 7.51 (dd, *J* = 2.5, 7.5 Hz, 2H, Ar-H), 7.67 (s, 1H, Ar-H), 8.39 (s, 1H, Ar-H); ¹³C NMR (125 MHz, CDCl₃): δ = 14.1 (C-7'), 22.6, 25.9, 29.0, 29.2, 31.7 (C-2', C-3', C-4', C-5', C-6'), 63.7 (C-5), 68.5 (C-1'), 68.8 (C-6), 72.9 (OCH₂Ph), 73.9 (C-3), 78.8 (C-2), 79.2 (C-4), 97.2 (C-1), 101.8 (CHPh), 110.0, 126.1, 127.6, 127.8, 128.2, 128.4, 129.1, 131.0, 137.2, 137.7 (15C, Ar-C), 183.6 (CS); HR-MS: *m/z*: Calcd for C₃₁H₃₈N₂NaO₆S [M+Na]⁺: 589.2348, found: 589.2351.

***n*-Heptyl 3-O-benzyl-4,6-O-benzylidene-2-deoxy- α -D-mannopyranoside (32).** A solution of **31** (238 mg, 0.419 mmol) in dry toluene (2 mL) was added dropwise over 10 min to a stirred solution of tributylstanne (0.169 mL, 0.63 mmol) in refluxing toluene (6 mL) under argon. After the reaction mixture was refluxed for 6 h, the solvent was removed *in vacuo* and the residue was purified by MPLC on silica (petroleum ether/EtOAc, 16:1-4:1) to afford **32** (107 mg, 58%) as a colorless oil. [α]_D +49.9 (c 1.07, MeOH); ¹H NMR (500 MHz, CDCl₃): δ = 0.89 (t, *J* = 7.0 Hz, 3H, H-7'), 1.30 (m, 8H, H-3', H-4', H-5', H-6'), 1.58 (m, 2H, H-2'), 1.81 (ddd, *J* = 3.5, 11.5, 13.5 Hz, 1H, H-2a), 2.28 (ddd, *J* = 1.0, 5.0, 13.5 Hz, 12H, H-2e), 3.35 (dt, *J* = 6.5, 9.5 Hz, 1H, H-1'a), 3.62 (dt, *J* = 6.5, 9.5 Hz, 1H, H-1'b), 3.70 (t, *J* = 9.0 Hz, 1H, H-4), 3.77 (t, *J* = 10.0 Hz, 1H, H-6a), 3.83 (m, 1H, H-5), 4.05 (ddd, *J* = 5.0, 9.0, 11.0 Hz, 1H, H-3), 4.26 (dd, *J* = 4.5, 10.0 Hz, 1H, H-6b), 4.69, 4.86 (d, *J* = 12.0 Hz, 2H, OCH₂Ph), 4.90 (d, *J* = 3.5 Hz, 1H, H-1), 5.63 (s, 1H, CHPh), 7.25-7.41 (m, 9H, Ar-H), 7.52 (dd, *J* = 2.5, 7.5 Hz, 1H, Ar-H); ¹³C NMR (125 MHz, CDCl₃): δ = 14.1 (C-7'), 22.6, 26.1, 29.1, 29.5, 31.8 (C-2', C-3', C-4', C-5', C-6'), 36.6 (C-2), 62.9 (C-5), 67.7 (C-1'), 69.2 (C-6), 73.0 (CH₂Ph), 73.1 (C-3), 84.0 (C-4), 97.9 (C-1), 101.3 (CHPh), 126.0, 127.5, 127.6, 128.2, 128.3, 128.8, 137.6, 138.8 (12C, Ar-C); HR-MS: *m/z*: Calcd for C₂₇H₃₆NaO₅ [M+Na]⁺: 463.2460, found: 463.2453.

***n*-Heptyl 2-deoxy- α -D-mannopyranoside (12).** A solution of **32** (66 mg, 0.15 mmol) in 80% aq. HOAc (1.25 mL) was stirred at 80 °C for 1 h and then concentrated to dryness. The residue was

purified by MPLC on silica (petroleum ether/EtOAc, 4:1-3:2) to afford the 3-O-benzylated intermediate. Hydrogenolysis of the intermediate was performed in MeOH (5 mL) in the presence of Pd(OH)₂/C (10%, 5.2 mg) under hydrogen (4 bar) at rt for 5 h. The reaction suspension was filtered through celite and the filtrate was concentrated *in vacuo*. The residue was purified by MPLC on silica (DCM/MeOH, 10:1) to provide **12** (18 mg, 46% over two steps) as a white solid. ¹H NMR (500 MHz, CD₃OD,): δ = 0.91 (t, *J* = 7.0 Hz, 3H, H-7'), 1.32 (m, 8H, H-3', H-4', H-5', H-6'), 1.57-1.62 (m, 3H, H-2e, H-2'), 2.04 (dd, *J* = 5.0, 13.0 Hz, 1H, H-2a), 3.23 (t, *J* = 9.5 Hz, 1H, H-4), 3.33 (m, 1H, H-1'a), 3.52 (m, 1H, H-5), 3.66-3.70 (m, 2H, H-6a, H-1'b), 3.79-3.85 (m, 2H, H-3, H-6b), 4.87 (d, *J* = 2.5 Hz, 1H, H-1); ¹³C NMR (125 MHz, CD₃OD,): δ = 14.4 (C-7'), 23.7, 27.4, 30.3, 30.7, 33.0 (C-2', C-3', C-4', C-5', C-6'), 39.0 (C-2), 62.9 (C-6), 68.3 (C-1'), 70.0 (C-3), 73.3 (C-4), 74.0 (C-5), 98.6 (C-1); HR-MS: *m/z*: Calcd for C₁₃H₂₆NaO₅ [M+Na]⁺: 285.1678, found: 285.1678.

Table S1. Statistics on X-ray diffraction data and refinement of FimH-CRD ligand complexes.

	FimH-CRD/2	FimH-CRD/5	FimH-CRD/6	FimH-CRD/7
Wavelength (Å)	1.00003	1.00000	0.97934	0.8000
Resolution range (Å)	24.69 - 2.00 (2.04 - 2.00)*	28.08 - 1.12 (1.16 - 1.12) *	42.15 - 1.65 (1.709 - 1.65) *	38.12 - 0.997 (1.032 - 0.997) *
Space group	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21
Unit cell	63.3 69.3 95.9	48.8 56.2 61.5	53.3 96.9 97.9	48.8 55.9 61
α, β, γ (°)	90 90 90	90 90 90	90 90 90	90 90 90
Unique reflections	50818 (2906)	60363 (4211)	57603 (6056)	90597 (8175)
Multiplicity	2.1 (2.0)	6.2 (4.6)	6.4 (6.3)	12.8 (12.4)
Completeness (%)	93.1 (99.6)	92.42 (65.37)	93.27 (99.46)	98.87 (89.84)
Mean I/sigma(I)	9.7 (3.9)	39.9 (7.1)	12.0 (2.1)	22.80 (2.06)
Wilson B-factor	13.27	10.95	14.48	10.53
R-meas	0.022 (0.21)	0.022 (0.21)	0.142 (1.05)	0.051 (0.97)
CC1/2	0.99 (0.99)	0.99 (0.95)	0.99 (0.69)	0.99 (0.79)
R-work	0.1745 (0.2418)	0.1077 (0.1201)	0.1696 (0.2639)	0.1236 (0.2280)
R-free†	0.2016 (0.2544)	0.1303 (0.1604)	0.2014 (0.2980)	0.1406 (0.2361)
Number of atoms	2965	1625	4884	1631
macromolecules	2447	1291	3666	1294
ligands	24	30	104	31
water	518	300	1108	298
Protein residues	316	158	474	158
RMS(bonds)	0.004	0.012	0.004	0.014
RMS(angles)	1.03	1.56	1.12	1.59
Ramachandran favored (%)	96.9	99.0	97.1	99.0
Ramachandran outliers (%)	0	0	0	0
Clashscore	0.21	2.31	2.15	3.81
Average B-factor	17.8	16	19.9	16.4
macromolecules	16.5	12.8	15.4	12.1
ligands	18.4	16.1	19.7	16.7
solvent	32.4	29.4	34.9	34.7

* Values in parentheses are for highest resolution shell.

In the refinement statistics, Friedel mates were merged. $R_S = \frac{\sum_h \sum_i |I_i(h) - \langle I(h) \rangle|}{\sum_h \sum_i I_i(h)}$, where i are the independent observations of reflection h .

† The R_{free} factor was calculated from 5% of the data, which were removed at random before the refinement.

Table S2. Thermodynamic quantities from ITC for binding of α -D-mannosides to FimH-CRD. All experiments were carried out at a temperature of 298.15 K and in 10mM HEPES buffer pH 7.4 containing 150mM NaCl. In case of competitive titrations, the concentration of the substance of interest and of the competitor **12** (see above for structure and synthesis) is given.

Ligand	Titration	Ligand [μ M] (strong/weak)	Protein [μ M]	K_D [nM]	ΔG° [kJ/mol]	ΔH° [kJ/mol]	$-T\Delta S^\circ$ [kJ/mol]	N	c-value
1	direct	100	10.0	29.1 (25.7 - 32.9)	-43.0	-49.8 (-49.4 - -50.3)	6.8	0.98	343
1	direct	100	8.6	29.0 (26.7 - 31.4)	-43.0	-50.9 (-50.5 - -51.2)	7.8	1.01	296
1				28.9 (25.8 - 32.3)	-43.0	-50.3 (-50.2 - -50.7)	7.3	1.00 \pm 0.02	320 \pm 34
2	direct	150	14.0	14.8 (13.9 - 15.9)	-44.7	-43.9 (-43.8 - -44.0)	-0.8	1.09	943
2	direct	100	10.5	21.1 (16.7 - 25.9)	-43.8	-46.2 (-45.6 - -46.8)	2.4	1.06	500
2				17.7 (14.1 - 22.3)	-44.2	-45.0 (-44.5 - -45.6)	0.8	1.07 \pm 0.02	722 \pm 313
3	direct	150	9.7	4.9 (4.2 - 5.8)	-47.4	-56.2 (-55.9 - -56.6)	8.8	1.05	1984
3	competitive	300 / 200 (13)	25.5	3.5 (2.7 - 4.6)	-48.3	-50.7 (-49.4 - -52.2)	2.4	0.99	378
3				3.5 (2.7 - 4.6)	-48.3	-56.2 (-55.9 - -56.6)	8.0	1.02 \pm 0.04	378
4	direct	100	10.0	2.2 (1.8 - 2.8)	-49.4	-60.9 (-60.5 - -61.3)	11.5	1.02	4448
4	competitive	300 / 300 (13)	32.1	1.3 (1.1 - 1.5)	-50.7	-57.3 (-56.6 - -58.2)	6.6	1.00	437
4				1.3 (1.1 - 1.5)	-50.7	-60.9 (-60.5 - -61.3)	10.1	1.01 \pm 0.01	437
5	direct	50	5.0	1.0 (0.7 - 1.3)	-51.4	-62.1 (-61.5 - -62.7)	10.7	1.00	10268
5	competitive	50 / 50 (13)	5.5	1.0 (0.7 - 1.4)	-51.3	-60.4 (-59.4 - -61.4)	9.1	0.99	990
5				1.0 (0.7 - 1.4)	-51.3	-62.1 (-61.5 - -62.7)	10.8	0.99 \pm 0.01	990
6	direct	110	8.7	14.0 (12.7 - 15.4)	-44.8	-63.6 (-63.2 - -64.0)	18.8	1.07	621
7	direct	38	4.4	7.2 (6.1 - 8.4)	-46.5	-72.6 (-71.8 - -73.4)	26.1	1.03	612
7	direct	80.75	9.8	4.9 (3.6 - 6.7)	-47.4	-71.1 (-70.1 - -72.2)	23.7	1.01	1984
7				6.2 (5.1 - 7.6)	-46.8	-71.8 (-71.0 - -72.7)	25.0	1.02 \pm 0.01	1298 \pm 970
8	direct	450	29.3	1'384 (1'206 - 1'553)	-33.4	-37.2 (-36.5 - -37.9)	3.7	1.00	21
8	direct	450	29.1	1'075 (980 - 1'179)	-34.1	-37.3 (-36.9 - -37.7)	3.2	0.95	27
8				1'222 (1'064 - 1'403)	-33.8	-37.2 (-36.4 - -38.0)	3.5	0.98 \pm 0.04	24 \pm 4
9	direct	170	11.3	144 (127 - 163)	-39.1	-39.1 (-38.6 - -39.7)	0.1	1.05	78
10	direct	400	27.0	1125 (980 - 1292)	-34.0	-42.9 (-42.1 - -43.8)	9.0	1.03	24
11	direct	4000	47.9	87.72 (82.81 - 92.92)	-23.2	-20.8 (-20.1 - -21.5)	-2.4	1.00*	0.5
11	direct	4000	42.1	92.11 (87.16 - 97.34)	-23.0	-21.4 (-20.7 - -22.1)	-1.7	1.00*	0.5
11				89.99 (86.14 - 94.02)	-23.1	-21.1 (-20.8 - -21.6)	-2.0	1.00*	0.5 \pm 0
12	direct	1000	50.4	9'330 (8'809 - 9'863)	-28.7	-19.9 (-19.5 - -20.4)	-8.8	1.00	5.4
12	direct	1000	50.1	9'456 (9'030 - 9'902)	-28.7	-19.1 (-18.9 - -19.4)	-9.5	1.01	5.3
12				9'386 (8'555 - 10'287)	-28.7	-19.5 (-19.1 - -20.0)	-9.1	1.00 \pm 0.00	5.3 \pm 0.1

* Stoichiometry was fixed to 1.00

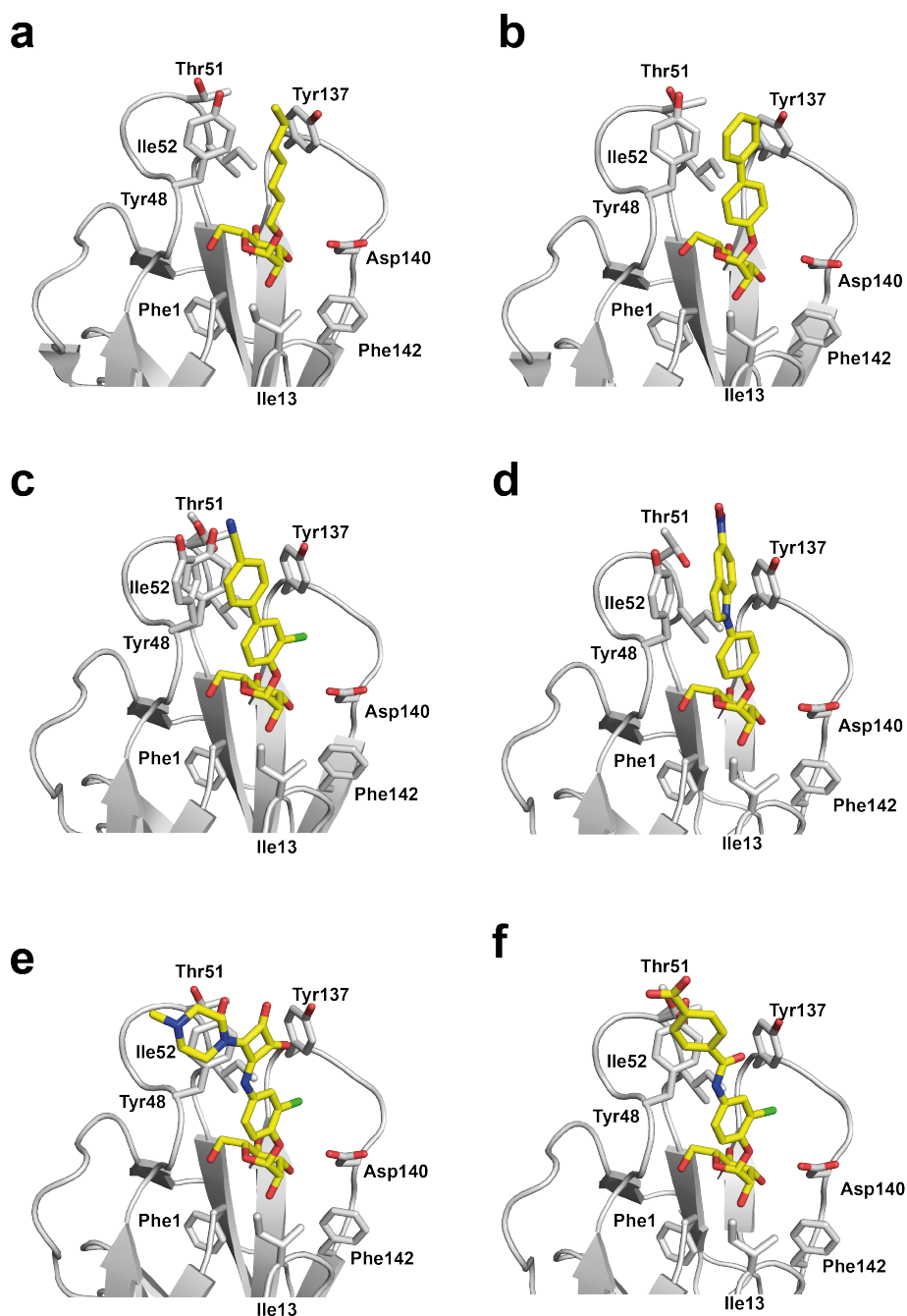


Figure S1. Comparison of the FimH antagonist binding site. The Mannose binding pocket of FimH cocrystal structures FimH-CRD/1 (PDB ID: 4BUQ) (a), FimH-CRD/2 (b), FimH-CRD/4 (PDB ID: 4CST) (c), FimH-CRD/5 (d), FimH-CRD/6 (e) and FimH-CRD/7 (f). The antagonists (yellow) and residues of the binding site (grey) are shown in stick representation.

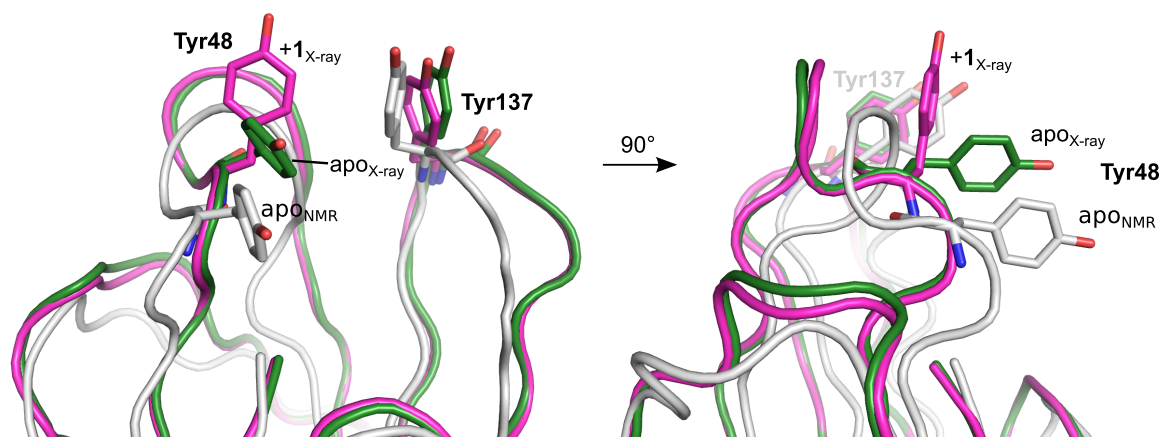


Figure S2. Trosine gate conformation in FimH-CRD X-ray and NMR solution structures. Superimposition of the “pseudo-apo” crystal structure of FimH-CRD with ethanediol in the binding site (green, PDB ID: 4AUU), the co-crystal structure with **1** (magenta, ligand not shown, PDB ID: 4BUQ) and the solution NMR structure of apo FimH-CRD published previously (grey, PDB ID: 3ZPD).^[S3] The conformations of the Tyr48 sides chain are similar in the two apo structures while in presence of antagonist the Tyr48 ring is flipped towards the binding loop. Tyr137 has a similar orientation in all structures.

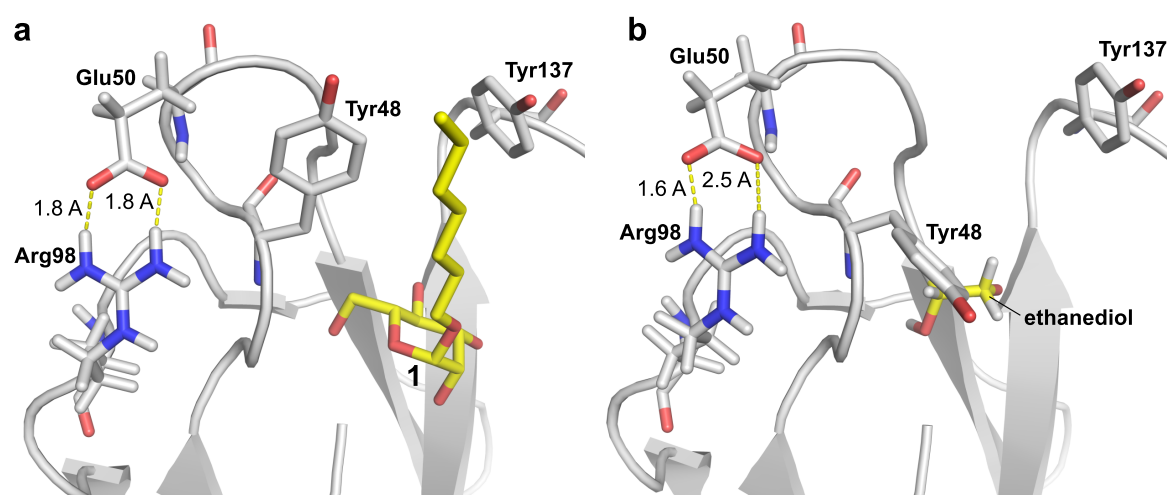


Figure S3. The side chain of Glu50 is stabilized by Arg98. FimH-CRD X-ray structures co-crystallized with **1** (PDB ID: 4BUQ) (a) and as "pseudo-apo" protein with ethanediol in the binding pocket (PDB ID: 4AUU) (b). The Glu50 side chain is stabilized by hydrogen bonds with the guanidinium group of Arg98 in both structures. The orientation of Glu50 and Arg98 in the co-crystal structures with antagonists **2**, **4**, **5**, **6** and **7** is similar to that with **1**.

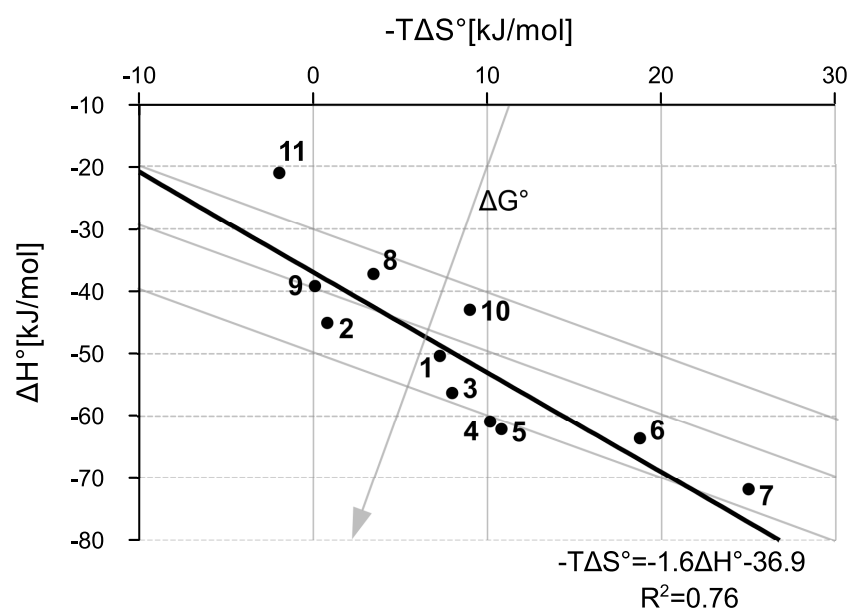


Figure S4. Enthalpy-entropy compensation of FimH antagonists. An enthalpy-entropy plot of FimH antagonists 1 to 11 obtained from ITC analysis is shown. A linear regression of all points (black line) suggests an average enthalpy gain of 1.6 kJ/mol for an entropy penalty of 1 kJ/mol. Grey diagonals indicate combinations of ΔH° and $-T\Delta S^\circ$ values resulting in the same ΔG° value. The grey arrow indicates increasing ΔG° values.

References

- [S1] a) A. Vervoort, C. K. De Bruyne, *Carbohydr Res* **1970**, 12, 277-280; b) J. De Prijcker, A. De Bock, C. K. De Bruyne, *Carbohydr Res* **1978**, 60, 141-53.
- [S2] S. G. Withers, D. J. MacLennan, I. P. Street, *Carbohydr Res* **1986**, 154, 127-144.
- [S3] S. Vanwetswinkel, A. N. Volkov, Y. G. Sterckx, A. Garcia-Pino, L. Buts, W. F. Vranken, J. Bouckaert, R. Roy, L. Wyns, N. A. van Nuland, *J Med Chem* **2014**, 57, 1416-1427.

Chapter 3.4. – Manuscript 6

Structural and Thermodynamic Characterization of the Interaction between the Bacterial Adhesin FimH and Mannosides: The Importance of the Tyrosine Gate

Preliminary manuscript formatted according to the requirements of *Journal of Biological Chemistry*.

Contributions

- | | |
|------------------------|--|
| S. Rabbani | <ul style="list-style-type: none">• Protein expression and purification.• Stability and competitive binding assays.• Manuscript preparation. |
| J. Bouckaert | <ul style="list-style-type: none">• Protein-ligand co-crystallization and data analysis• Manuscript preparation |
| R. C. Preston | <ul style="list-style-type: none">• Isothermal titration calorimetry.• Manuscript preparation (thermodynamic part). |
| A. Zalewski and S. Eid | <ul style="list-style-type: none">• Molecular Modeling.• Manuscript preparation (modeling part). |
| A. Thompson | <ul style="list-style-type: none">• X-ray diffraction measurements |

Structural and Thermodynamic Characterization of the Interaction between the Bacterial Adhesin FimH and Mannosides: The Importance of the Tyrosine Gate

(Short title: FimH tyrosine gate)

Said Rabbani¹, Julie Bouckaert², Adam Zalewski¹, Roland Preston¹, Sameh Eid¹, Andrew Thompson³,
and Beat Ernst^{1*}

¹) Institute of Molecular Pharmacy, Pharmacenter, University of Basel, Switzerland
Klingelbergstrasse 50-70, CH-4056 Basel, Switzerland

²) Unité de Glycobiologie Structurale et Fonctionnelle, UMR 8576 du CNRS, Villeneuve d'Ascq,
France

³) Synchrotron SOLEIL, L'Orme de Merisiers, Saint Aubin-BP48, Gif-sur-Yvette Cedex, France.

^{}) To whom correspondence should be addressed*

Beat Ernst

Institute of Molecular Pharmacy

Pharmacenter

University of Basel

Klingelbergstrasse 50-70

CH-4056, Basel

Switzerland

Fax +41 61 267 15 52

E-mail: beat.ernst@unibas.ch

Abbreviations

ABTS, 2,2'-azino-di-(3-ethylbenzthiazoline-6-sulfonic acid); BSA, bovine serum albumin; CRD, carbohydrate recognition domain; DMSO, dimethyl sulfoxide; EDTA, ethylenediaminetetraacetic acid; FCS, fetal calf serum; IPTG, isopropyl 1-thio- β -D-galactopyranoside; ITC, isothermal titration calorimetry; MM-GBSA, molecular mechanics-generalized born surface area; PAA, polyacrylamide; HEPES, 4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid; TM-PAA, Man α 1-3(Man α 1-6)Man β 1-4GlcNAc β 1-4GlcNAc β -PAA; UPEC, uropathogenic *Escherichia coli*; and UTI, urinary tract infection.

Keywords

Uropathogenic *Escherichia coli*, urinary tract infection, bacterial adhesin FimH, FimH antagonists, tyrosine gate, isothermal titration calorimetry, competitive binding assay, molecular dynamics simulations.

Summary

Urinary tract infections (UTIs) are among the most prevalent infections worldwide resulting in considerable financial burden on society. Women have a 40-50% risk to experience at least one symptomatic UTI episode during their life. The disease is most frequently caused by uropathogenic *Escherichia coli* (UPEC). The type 1 pili tip FimH allows UPEC to bind specifically to the high-mannosylated uroplakin-Ia glycoprotein on the bladder epithelium and thereby enables the invasion in the host cells. High affinity mannoside-based FimH antagonists that block the adhesion of UPEC are therefore considered as attractive anti-infective therapeutics. Alkyl- and aryl-functionalized FimH antagonists bind in the *in-docking* or *out-docking* mode respectively, as revealed by X-ray studies and supported by computer-based docking calculations. In the *in-docking* mode the alkyl-aglycone of the mannoside is hosted by two tyrosine residues (Y48 and Y137) located at the entrance of the binding site. In the *out-docking* mode, the bulky biphenyl aglycones form π - π interactions exclusively with the Y48 side chain, which account for higher affinity. To investigate the contribution of both tyrosine residues, the binding affinities and thermodynamic profiles of an alkyl- and an aryl substituted mannoside were assessed on wild type FimH-carbohydrate recognition domain (CRD) and the corresponding Y48A and Y137A mutants. Co-crystal structures of the three proteins with heptyl α -D-mannopyranoside were also obtained.

Different contributions of the tyrosine gate residues were observed, however not in a ligand-dependent manner. In a competitive binding assay, the Y48A mutant exhibited similar binding affinities as the wild type for the alkyl- and aryl mannosides, whereas binding was reduced 6- to 8-fold for the Y137A mutant. Thermodynamic profiling showed significant entropy-enthalpy compensation for the Y48A mutant, whereas only a partial compensation was observed for the Y137A mutant. Molecular modeling revealed a reorganization of the binding site upon Y137A mutation. The thermodynamic and structural data presented here improve our understanding of FimH-mannoside interactions.

Introduction

Urinary tract infections (UTIs) are the most common bacterial infections in human (1, 2). They are mainly caused by type 1-piliated uropathogenic *Escherichia coli* (UPEC) (3). The pili are proteinaceous hair-like appendages composed of four different subunits (FimH, FimG, FimF and FimA), which are assembled in the periplasmic space with the assistance of the chaperon-like FimC subunit via the conserved Usher pathway (4). The FimH subunit is located at the tip of the pili (5) and consists of two structural domains: (i) the N-terminal carbohydrate recognition domain (CRD), which binds specifically to mannose moieties displayed on the uroplakin complexes at the luminal surface of bladder epithelial cells (6-8), (ii) the pilin domain located at the C-terminus connects FimH to the core of the pili via FimG subunit. The attachment of bacteria to the bladder epithelium is followed by their internalization and subsequent biofilm formation inside the urothelial cells (9-11), which facilitate the persistence of UPEC within the urinary tract. FimH-mediated binding is enhanced under shear stress (12) whereas under static conditions the interactions between lectin and pilin domains increase at the expense of affinity towards the ligand (13).

Antibiotics are the standard treatment for UTIs (14-16). However, their long-term use promotes the development of microbial resistance leading to recurrent infections (17,18), and thus accounting for significant morbidity and high medical costs (19). Therefore, anti-adhesion drugs, which ensure the clearance of UPEC from the urinary tract, exhibit a promising therapeutic potential as anti-infectives (20-28). In addition, such therapeutics are less likely to induce bacterial resistance, because their effect is neither bacteriostatic nor bacteriolytic.

For the rational design of FimH antagonists, several 3D-structures are available. The first crystallographic structure of FimH was in complex with its chaperone FimC (29). The mannose-binding site was occupied by one molecule of cyclohexylbutanoyl-*N*-hydroxyethyl-D-glucamide (C-HEGA), an additive for the crystallization. The second structure of FimH-FimC bound to α -D-mannopyranose was reported in 2002 (30). The ligand was located in a deep pocket in which the residues F1, N46, D47, D54, Q133, N135, D140 and F142 create an optimal hydrogen network with all the hydroxyl groups of α -D-mannopyranose, with the exception of the anomeric hydroxyl group. At the entrance of the mannose-binding pocket, the side chains of the tyrosine residues 48 and 137, in an almost parallel orientation, build the so-called “tyrosine gate”. In another crystallographic study a *n*-butyl α -D-mannopyranoside molecule originating from the culture medium (LB-medium) was found within FimH-CRD (31). The butyl aglcone was accommodated in a hydrophobic environment by van der Waals contacts within the tyrosine gate, with further participation of I52. The structure of FimH-CRD in complex with oligomannose-3 (32) illustrated the crucial role of the tyrosine gate for natural ligand binding, as the two residues allow for contacts to the hydrophobic face of the carbohydrates. Therefore, the binding-mode in which portions of the molecule are located within the

tyrosine 48 and tyrosine 137 side chains are termed the *in-docking* mode. In contrast, an *out-docking* mode where the biphenyl aglycone is not accommodated within the tyrosine-gate and forms interactions to Y48 exclusively was observed (23). In both binding modes favorable hydrophobic/ π - π stacking interactions with Y48 account for higher ligand affinity (23, 31-33).

To establish a detailed interaction profile and to rationalize binding affinities of various FimH antagonists we generated the tyrosine gate alanine mutants (Y48A and Y137A). In the first part of this project, we determined the experimental affinities and the thermodynamic binding profiles of selected mannosides. In the second part, we investigated the binding characteristics of FimH *wt* and mutants focusing on structures co-crystallized with *n*-heptyl α -D-mannopyranoside. The thermodynamic and structural data presented here improve considerably our understanding of FimH-mannoside interaction and guide for better design strategy of FimH antagonists.

Material and Methods

Reagents

Bacto-Yeast extract, Bacto-Agar, and Bacto-Tryptone were purchased from Becton Dickinson (Basel, Switzerland) and used for the preparation of LB (Luria-Bertani) culture medium. Isopropyl 1-thio- β -D-galactopyranoside (IPTG) was obtained from Applichem (Darmstadt, Germany). Polymyxin B sulfate, Hepes (4-(2-hydroxyethyl)-piperazine-1-ethanesulfonic acid), oxalic acid, MgCl_2 , CaCl_2 , NaH_2PO_4 , imidazole were purchased from Fluka (Buchs, Switzerland). Ampicillin, bovine serum albumin (BSA) and ethylenediaminetetraacetic acid (EDTA), amino acids and BME vitamin mix were obtained from Sigma (Buchs, Switzerland). 1,5-anhydro D-mannitol (**1**), *n*-heptyl α -D-mannopyranoside (**2**) and 4-biphenyl α -D-mannopyranoside **3** were synthesized as described in (34, 35 and 27). The biotinylated polyacrylamide (PAA) glycopolymer $\text{Man}\alpha 1\text{-}3(\text{Man}\alpha 1\text{-}6)\text{Man}\beta 1\text{-}4\text{GlcNAc}\beta 1\text{-}4\text{GlcNAc}\beta\text{-PAA-biotin}$ (TM-PAA) containing 20 mol% sugar residues and 5 mol% biotin was purchased from Lectinity (Moscow, Russia). MaxiSorp 96-well microtiter plates were purchased from Nunc (Roskilde, Denmark).

Cloning of tyrosine gate mutants

FimH-CRD construct linked to the thrombin cleavage site (Th), and a 6His-tag (6His) was generated as described (36). The Y48A and AY137 mutants were generated by overlap extension PCR method (37) using the wild type (*wt*) encoding plasmid as template. The inserts were digested with HindIII and XbaI restriction enzymes, gel purified and subsequently ligated into the corresponding cloning site of pDsbA3 expression vector (36). The vectors were then transformed into *E. coli* DH5 α chemo-competent cells (Novagen, Lucerne, Switzerland). After plasmid isolation and restriction control, the correctness of the constructs was confirmed by DNA sequencing (Microsynth, Balgach, Switzerland). Finally, the vectors were transformed into the protease-deficient *E. coli* strain HM 125 (38) for protein expression. For protein crystallization a FimH-CRD construct (amino acids: 1-158), without any tag, was generated and purified by ion exchange chromatography as described (39).

Protein expression and purification

Bacterial clones were grown at 30 °C with vigorous shaking (300 rpm) in M9 minimal medium (40) supplemented with 2 mM MgSO_4 , 0.1 mM CaCl_2 , 20 % glucose, 10 μM of each amino acid, BME vitamin mix (Sigma, Buchs, Switzerland) and 100 $\mu\text{g/mL}$ ampicillin. When an OD_{600} of 0.8 was reached, the cells were induced with 1 mM IPTG and further cultivated for 16 h at 30 °C and 300

rpm. Then the cells were cooled on ice for 5 min and harvested by centrifugation at 5,000 rpm for 20 min at 4 °C. The pellet was suspended in a cold solution of 50 mM Tris/HCl, pH 7.5, 150 mM NaCl, 5 mM EDTA, 1 mg/mL polymyxin B sulfate (lysis buffer) and stirred for 2 h at 4 °C. After centrifugation at 11,000 rpm for 20 min at 4 °C, the supernatant (periplasmic extract) was dialyzed overnight against 50 mM NaH₂PO₄, 300 mM NaCl, and 10 mM imidazole/NaOH, pH 8 (binding buffer), and applied to a Ni-NTA column (Sigma, Buchs, Switzerland) attached to a BioLogic fast protein liquid chromatography system (BioRad, Reinach BL, Switzerland). The column was washed with binding buffer and afterwards eluted with 50 mM NaH₂PO₄, 300 mM NaCl, and 250 mM imidazole/NaOH, pH 8 (elution buffer). The fractions containing FimH-CRD-Th-6His were pooled and dialyzed against water and then against 20 mM HEPES/NaOH, 150 mM NaCl, and 1 mM CaCl₂, pH 7.4 (assay buffer). The purity of the protein was verified by SDS–PAGE analysis, and the amount was determined by HPLC (41) using BSA as standard. Aliquots of the proteins could be stored for up to 3 months at 4 °C. For long-term storage, the proteins were frozen at -80°C without additives.

Circular dichroism analysis

The far-UV region of a protein circular dichroism (CD) spectrum delivers spectral features that are directly relevant to the secondary structure of the protein (42, 43). To analyze the effect of the mutations on the secondary structure of FimH-CRD-Th-6His CD-spectrum profiles of the wild type and mutants were recorded. The proteins were diluted to 10 µM concentration in 10 mM sodium phosphate buffer at pH 7.4 and CD spectra were measured at 25°C using thermostat-controlled 0.1 cm cell. The spectra were recorded between 190 and 250 nm with 2 nm bandwidth and 10 scans per sample on a Jasco J-720 spectrophotometer with temperature controller PFD-350S connected to the software program J-700. CD-spectra were corrected for background by subtracting buffer spectra and the CD signal was converted to residue ellipticity ($[\theta]_{MRW}/\text{deg.cm}^2 \text{ dmol}^{-1}$).

Fluorescence measurements

The conformational stability was verified by monitoring the GdmCl-induced unfolding profiles (44) of the wild type and mutants. Prior spectroscopic measurements the samples (10 µM) were incubated with GdmCl concentrations ranging from 0 to 6 M for 24 h at 25°C. The effective GdmCl concentrations were determined via the refractive index of the samples (45). Fluorescence measurements were performed on a Hitachi F4500 fluorescence spectrometer at 25°C using an excitation slit of 2.5 nm and an emission slit of 5 nm. All samples were excited at $\lambda = 280$ nm and the spectra were recorded at a scan speed of 10 nm/min. The fluorescence emission for equilibrium unfolding experiments was followed at $\lambda = 350$ nm (tryptophan fluorescence). Fluorescence spectra

were corrected for background by subtracting buffer spectra. All transitions were evaluated according to a two-state model using the linear extrapolation method and normalized.

Protein MS analysis

The exact molecular weight of FimH-CRD wild type and mutants was verified by matrix-assisted laser desorption ionization mass spectrometry (MALDI-MS). To differentiate between Y48A and Y137A mutants the proteins were subjected to trypsin digestion followed by MS analysis.

Competitive binding assay

To evaluate the affinity of the mutants a competitive binding assay described previously (36) was applied. Microtiter plates (F96 MaxiSorp, Nunc) were coated with 100 μ L/well of a 10 μ g/mL solution of FimH-CRD-Th-6His in 20 mM HEPES/NaOH, 150 mM NaCl and 1 mM CaCl_2 , pH 7.4 (assay buffer) overnight at 4 $^{\circ}\text{C}$. The coating solution was discarded and the wells were blocked with 150 μ L/well of 3% BSA in assay buffer for 2 h at 4 $^{\circ}\text{C}$. After three washing steps with assay buffer (150 μ L/well), a serial dilution of the test compound (50 μ L/well) in assay buffer containing 5% DMSO and streptavidin-peroxidase coupled TM-PAA polymer (50 μ L/well of a 0.5 μ g/mL solution) were added. The plates were incubated for 3 h at 25 $^{\circ}\text{C}$ and 350 rpm and then carefully washed four times with 150 μ L/well assay buffer. After the addition of 100 μ L/well of ABTS-substrate, the colorimetric reaction was allowed to develop for 4 min, then stopped by the addition of 2% aqueous oxalic acid before the optical density (OD) was measured at 415 nm on a microplate-reader (Spectramax 190, Molecular Devices, California, USA). The IC_{50} values of the compounds tested in duplicates were calculated with prism software (GraphPad Software, Inc., La Jolla, USA). The IC_{50} defines the molar concentration of the test compound that reduces the maximal specific binding of TM-PAA polymer to FimH-CRD by 50%. The relative IC_{50} (rIC_{50}) is the ratio of the IC_{50} of the test compound to the IC_{50} of *n*-heptyl α -D-mannopynoside. The relative IC_{50} (rIC_{50}) was set to 1 for the wild type.

Isothermal titration calorimetry

The thermodynamic characterization of the interactions between FimH-CRD wild type, Y48A and Y137A and mannosides was measured by isothermal titration calorimetry (ITC) as previously described (27). The protein was dialyzed over night at 4 $^{\circ}\text{C}$ against assay buffer using Slide-A-Lyzer dialysis cassettes with 10 kDa cut-off (Thermo Fisher Scientific, Waltham, MA, USA). Final protein concentrations of 8-15 μ M for the wild type and the Y48A mutant, and 17-23 μ M for the Y137A mutant were used. The ligands 1,5-anhydro D-mannitol (**1**), *n*-heptyl α -D-mannopyranoside (**2**), and

4-biphenyl α -D-mannopyranoside (**3**) were diluted to 15-25 fold molar excess of the protein. Dilution enthalpy was determined by ligand in assay buffer titration. For the 4-biphenyl α -D-mannopyranoside (**3**), a final concentration of 2.5 % DMSO was required to dissolve the compound. For these measurements, DMSO was added to the protein samples (2.5 % final concentration). All measurements were performed with a MicroCalTM VP-ITC instrument (GE Healthcare, Northampton, MA, USA; sample cell volume of 1.4523 mL) at 25 °C, 307 rpm stirring speed, and 10 μ cal/s reference power. The samples were preheated to 22 °C and degassed for 5 minutes prior to the measurements. Titration was started after steady baseline equilibration was reached with an initial 2 μ L injection, which was excluded from subsequent data analysis. Ligand was injected in 4-7 μ L steps (38-66 injections in total) with a spacing of 10 minutes to ensure non-overlapping peaks. Sigmoidal binding curves with complete saturation at the end of each experiment were obtained. Baseline adjustments and peak integration were performed using the Origin 7 software (OriginLab, Northampton, MA, USA). The parameters K_D (dissociation constant), N (stoichiometry), and ΔH° (change in enthalpy) were determined using non-linear least-square fitting in Microsoft Excel and the Solver add-in (Frontline Systems Inc., Incline Village, NV) according to Ziegler *et al.* (46). The thermodynamic parameters ΔG° (change in free energy) and $T\Delta S^\circ$ (change in enthalpy) were calculated using equation 1,

$$\Delta G^\circ = \Delta H^\circ - T\Delta S^\circ = RT\ln K_A = -RT\ln K_D \quad (\text{Eq. 1})$$

where T is the absolute temperature and R is the universal gas constant (8.314 J mol⁻¹ K⁻¹). The quantity $c = Mt(0) K_D^{-1}$, where $Mt(0)$ is the initial macromolecule concentration, ranged between 10 and 600 for all experiments, thus being in the reliable range (47).

Molecular modeling

For all experiments a 1.69 Å resolution crystal structure (PDB code 1UWF) of the FimH-CRD was used (31). Initial protein-ligand complexes were obtained through flexible docking using Glide (Glide, version 5.5, Schrödinger, LLC, New York, NY, 2009). Molecular dynamics (MD) simulations were carried out with Desmond (48) using the OPLS 2005 force field as implemented in the Schrödinger 2011 suite (Maestro, version 9.2, Schrödinger, LLC, New York, NY, 2011). Each system was solvated using an orthorhombic, TIP3P water (49) box with a minimum distance of 10 Å from the complex. Na⁺ and Cl⁻ ions were added to neutralize the charges and account for physiological salt concentration (0.15 M). Long-range electrostatic interactions were handled using the particle mesh Ewald summation (50). All systems were equilibrated using the default relaxation protocol (Desmond 2.2; Schrodinger, Inc., New York, NY) and simulated over the span 30 ns with a time step of 2.0 fs. The SHAKE algorithm (51) was applied to all heavy-atom bound hydrogens. Production runs were

carried out in the Martyna-Tobias-Klein isothermal-isobaric ensemble (NPT) (52) using the Nose-Hoover barostat to maintain a constant temperature of 300 K (53). Energetic and structural data were recorded in 3.0 ps intervals. Trajectory clustering was performed based on ligand heavy atom RMSDs (1000 snapshots per simulation with 0.5 Å cutoff between clusters) using a hierarchical average-linkage algorithm. Molecular mechanics-generalized Born surface area (MM-GBSA) calculations were performed on ensembles of 250 snapshots using the default settings implemented in Prime (Prime, version 3.0, Schrödinger, LLC, New York, NY, 2011). All images were generated using Maestro and PyMOL (54).

Crystallization, data collection and refinement

Crystallization of the FimH lectin was accomplished using the sitting drop vapor diffusion method (Hampton Research Crystal Screens I and II). Data were collected on Proxima I (SOLEIL, Saint-Aubin, France) and PX14.2 (PSF BESSY, Berlin, Germany). The structure determination was performed using molecular replacement by the program PHASERi with apo-FimH (PDB entry code 4AUU) as a model. The ligands were parametrised using the ProdrG server (55) and crystal structures were refined using Phenix (56). Data refinement and statistics are summarized in Table 6S (supplementary material).

Results

Tyrosine gate mutants are conformationally stable

For the treatment of UTIs, several FimH antagonists have been rationally designed, synthesized and evaluated (22-28). The interaction of the aryl aglycone of mannosides with the tyrosine gate (Y48 and Y137), have been shown to significantly improve the affinity. To experimentally evaluate the role of these residues we generated the corresponding alanine mutants (Y48A and Y137A). The proteins were expressed in the proteases deficient *E. coli* HM 125 and purified from the periplasmic space by chromatography on a Ni-NTA affinity matrix to homogeneity (Figure 1A). The mutations were confirmed at the protein level by ESI-MS (Figure S1).

The stability of the mutants was preserved as verified in GdmCl unfolding experiments (Figure 1B). The conformational transition curves (Figure 1B) of the wild type and the mutants show superimposable profiles with similar transition mid-point values $D_{1/2}$ (2.75 M GdmCl (wild type), 2.77 M GdmCl (Y48A) and 2.73 M GdmCl (Y137A)). The data were evaluated according to the two-state model yielding a free energy of folding $\Delta G_{H_2O}^\circ$ of -50.00 ± 3.07 kJ mol⁻¹ (wild type), -44.05 ± 2.37 kJ mol⁻¹ (Y48A) and -53.71 ± 4.53 kJ mol⁻¹ (Y137A) and a folding cooperativity of 18.17 ± 1.11 kJ mol⁻¹, 15.95 ± 0.85 kJ mol⁻¹ and 19.63 ± 1.64 kJ mol⁻¹ for the wild type, Y48A and Y137A mutants, respectively. To evaluate whether the secondary structure of the proteins was altered upon the mutation, CD-spectra of the Y48A and Y137A mutants were recorded and compared to the corresponding CD-spectrum of the wild type (Figure 1C). The result shows identical spectrum profiles indicating that the mutants in native conformation retain secondary structure as the wild type.

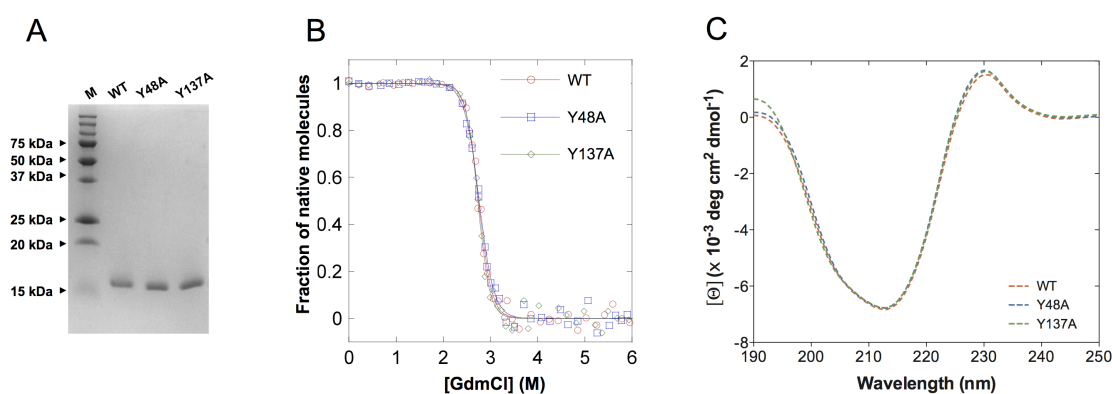


Figure 1. Characterization of FimH-CRD wild type and mutants: (A) Affinity-purified FimH-CRD wild type (WT) and tyrosine gate mutants Y48A and Y137A were subjected to SDS–PAGE under reducing conditions on a 16% gel. M, molecular weight marker. (B) GdmCl-dependent equilibrium unfolding profiles at 25 °C and pH 7.4 were monitored by changes in fluorescence at 350 nm upon excitation at 280 nm. All transitions were evaluated according to a two-state model using the linear extrapolation method and normalized, red line (WT), blue line (Y48A) and green line (Y137A). (C) CD-spectra of FimH-CRD wild type (red line), Y48A (blue line) and Y137A (green line). All samples were measured at 10 μM concentration in 10 mM sodium phosphate buffer pH 7.4 and at 25°C using thermostat-controlled 0.1 cm cell as described in material and methods section.

Contributions of Y48 and Y137 to ligand binding

The affinities of 1,5-anhydro D-mannitol (**1**), *n*-heptyl α -D-mannopyranoside (**2**) and the 4-biphenyl α -D-mannopyranoside (**3**) (Figure 2A) were evaluated with the wild type and the mutants in a competitive binding assay (36). The affinities are reported relative to **1** as rIC_{50} (Figure 2B). The mutation Y48A did not affect the binding properties of FimH-CRD, *i.e.* the rIC_{50} values for the tested mannosides were comparable to those obtained with the wild type. In contrast, the Y137A mutation resulted in a dramatic decrease in affinity towards *n*-heptyl α -D-mannopyranoside (**2**) and 4-biphenyl α -D-mannopyranoside (**3**), whereas the affinity to 1,5-anhydro D-mannitol (**1**) was only slightly decreased compared to the wild type. Compound **1** lacks the aglycone moiety and the anomeric oxygen and is therefore, not predicted to be affected by the mutations at the tyrosine gate. In wild type FimH-CRD, the aglycone of alkyl mannosides has been shown to interact with side chains of both tyrosine residues (Y48 and Y137, *in-docking*) (31), whereas in case of the bulky aromatic aglycone exclusively with Y48 (*out-docking*). It was therefore surprising that the Y137A mutation, and not Y48A mutation, was detrimental for the binding of 4-biphenyl α -D-mannopyranoside (**3**).

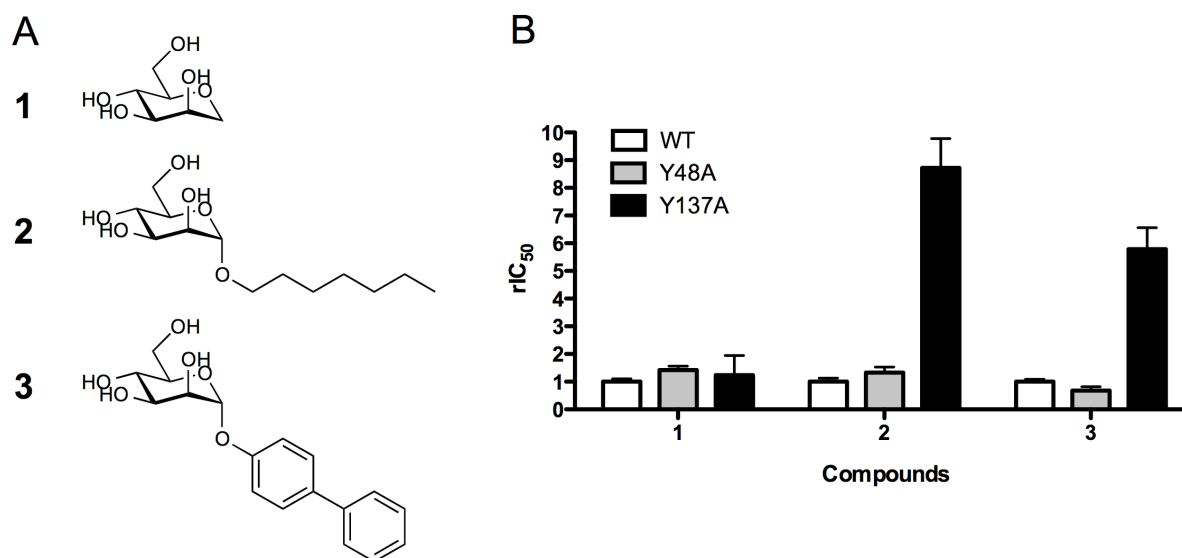


Figure 2. Structure and affinity of FimH ligands. (A) 1,5-anhydro D-mannitol (**1**), *n*-heptyl α -D-mannopyranoside (**2**), and 4-biphenyl α -D-mannopyranoside (**3**). (B) The relative affinity of FimH-CRD wild type, Y48A and Y137A towards compounds **1**, **2** and **3** was determined by the competitive binding assay (36).

Thermodynamic profiles of FimH mutants

To gain further insights into the interaction of the antagonists **1**, **2**, and **3** with each of the FimH variants (wild type, Y48A and Y137A), direct binding affinity as well as the thermodynamic fingerprints were determined. Isothermal titration calorimetry (ITC) was employed for the measurements as described in material and methods section. The results are shown in Table 1 and Figure 3. The binding of the three compounds (**1-3**) was largely enthalpy driven, which is in agreement with previously reported results for other mono- and oligomannoside derivatives (33, 57) and for a series of biphenyl-mannosides (27). For the mannoside derivative 1,5-anhydro D-mannitol (**1**) K_D values in the low micromolar range were obtained, which was expected since beneficial interactions between the aglycone moiety and the tyrosine-gate were missing. For this compound both mutants showed a 2.5 fold decrease in binding affinity compared to the wild type. This was due to a slightly decreased binding enthalpy, which was not entirely compensated by an increase in entropy. Such small effects might arise by slight modification of amino acid side chains or altered solvation properties. The *n*-heptyl α -D-mannopyranoside (**2**), expected to dock between the tyrosine gate residues, exhibited a K_D of 31.3 nM for the wild type, which is in agreement with reported values measured by surface plasmon resonance (31). The Y48A mutant showed a two-fold decreased affinity (61.9 nM). Probably, the removal of the Y48 residue resulted in a loss in enthalpy ($\Delta\Delta H = +6.5$ kJ mol⁻¹), likely due to missing hydrophobic interactions of the alkyl moiety with the phenyl side chain. The increased flexibility in the alkyl moiety was partially compensated by beneficial entropy ($-T\Delta\Delta S = -5.5$ kJ mol⁻¹), resulting in only a small net loss in binding free energy ($\Delta\Delta G = +1.0$ kJ mol⁻¹). The Y137A mutant exhibited a similar beneficial entropic effect ($-T\Delta\Delta S = -6.1$ kJ mol⁻¹) but suffered from a more pronounced loss in enthalpy ($\Delta\Delta H = +10.9$ kJ mol⁻¹), resulting in a seven-fold reduction of the affinity ($\Delta\Delta G = +4.8$ kJ mol⁻¹). The 4-biphenyl α -D-mannopyranoside (**3**) was expected to bind outside the tyrosine gate (23) and display little interaction with the residue Y137. This compound exhibited slightly improved overall binding affinity compared to compound **2**, with a K_D of 19.2 nM for FimH-CRD wild type. Similarly as observed for compounds **1** and **2**, 4-biphenyl α -D-mannopyranoside (**3**) showed a two-fold decreased affinity towards the Y48A mutant when compared to the wild type. The differences in enthalpy ($\Delta\Delta H = +4.4$ kJ mol⁻¹) and entropy ($-T\Delta\Delta S = -2.9$ kJ mol⁻¹) contributions were however, not as pronounced as for compound **2**. Considering the loss of π - π interactions with Y48 side chain, these changes are surprisingly small. Interestingly, the binding to Y137A showed a significant loss in enthalpy ($\Delta\Delta H = +12.1$ kJ mol⁻¹). This detrimental enthalpy was partially compensated by a beneficial entropy contribution ($-T\Delta\Delta S = -8.4$ kJ mol⁻¹). While for compound **2**, the entropy term was similar for Y48A and Y137A mutants (-8.4 kJ mol⁻¹ vs. -9.0 kJ mol⁻¹, respectively) there was a substantial gap for compound **3** (-2.5 kJ mol⁻¹ vs. -8.0 kJ mol⁻¹). This

suggests that the biphenyl-FimH interface retains greater mobility for the Y137A mutant, whereas compound **2** exhibits similar mobility regardless of the mutated tyrosine residue.

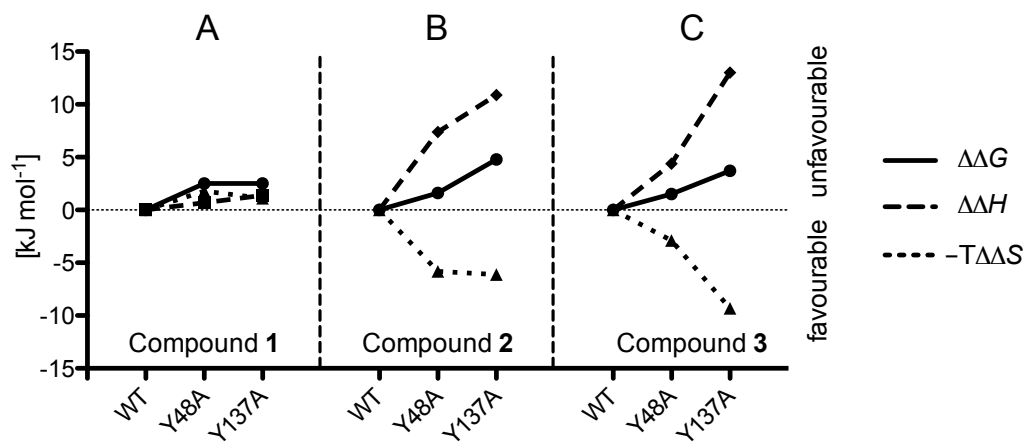


Figure 3. The thermodynamic fingerprint of the binding of 1,5-anhydro D-mannitol (**1**), *n*-heptyl α -D-mannopyranoside (**2**), and 4-biphenyl α -D-mannopyranoside (**3**) to FimH-CRD wild type and mutants (Y48A & Y137A). The parameters $\Delta\Delta G$, $\Delta\Delta H$ and $-T\Delta\Delta S$ are plotted relative to compound **1**.

Table 1. Thermodynamic parameters of compounds **1**, **2**, and **3** binding to FimH-CRD wild type, Y48A, and Y137A mutants were determined by isothermal titration calorimetry. The dissociation constant (K_D), the change in enthalpy (ΔH), and the molar ratio of protein and ligand (N) were determined from the non-linear least squares fitting as described in the material and methods section. The change in free binding energy (ΔG) and change in entropy ($-T\Delta S$) were calculated with equation 1 ($\Delta G = \Delta H - T\Delta S$).

Compound	Protein	N	rK_D	K_D [nM]	ΔG° [kJ mol $^{-1}$]	ΔH° [kJ mol $^{-1}$]	$-T\Delta S^\circ$ [kJ mol $^{-1}$]
1,5-Anhydro D-mannitol (1)	WT	1.07	1	943	-34.5	-36.6	2.1
	Y48A	1.05	2.6	2490	-32.0	-35.9	3.9
	Y137A	1.06	2.6	2476	-32.0	-35.2	3.2
<i>n</i>-Heptyl α-D-mannopyranoside (2)	WT	1.00	1	31.3	-42.8	-39.9	-2.9
	Y48A	0.98	2.0	61.9	-41.8	-33.4	-8.4
	Y137A	1.05	7.1	220.7	-38.0	-29.0	-9.0
4-Biphenyl α-D-mannopyranoside (3)^a	WT	1.00	1	19.2	-44.0	-44.4	0.4
	Y48A	1.05	1.8	35.5	-42.5	-40.0	-2.5
	Y137A	1.04	4.5	87.3	-40.3	-32.3	-8.0

^a Measurement performed with 2.5 % DMSO for solubility reasons

Molecular modeling

In both competitive binding assay and ITC the Y48A mutation has shown only a moderate effect on the affinity of FimH-CRD towards *n*-heptyl α -D-mannopynoside (**2**) and 4-biphenyl α -D-mannopyranoside (**3**). This result was not expected, since the interaction between the mannoside aglycones, in both the *in-docking* and the *out-docking*, and the side chain of Y48 was no longer possible. Molecular modeling simulations were applied to elucidate the underlying reasons for this unexpected result. MM-GBSA binding free energy calculations performed on ensembles of ligand-bound MD snapshots showed that the mutation Y48A resulted in diminished protein-ligand interaction energies (Tables S1 and S2). At the same time, visual inspection and clustering analyses of the simulations indicated that the removal of the bulky side chain allowed for increased ligand aglycone mobilities, especially in the case of *n*-heptyl α -D-mannopyranoside (**2**) (Figure 4A and 4C). Furthermore, given that the Y48 side chain is, upon binding, forced to adopt a single, mode-dependent rotamer, a corresponding increase in entropic costs are expected. In summary, these opposing effects result in an overall improvement of the entropy term. The flexibility of 4-biphenyl α -D-mannopyranoside (**3**) however, remains similar (Figure 4B and 4D). This agrees with the thermodynamic profile of the binding of antagonist **3**, which did not show an as pronounced enthalpy-entropy compensation as observed for *n*-heptyl α -D-mannopynoside (**2**) with the Y48A mutant.

Additional MD simulations for the *apo* wild type and mutant proteins coupled with detailed geometric and energetic analysis indicated significant deformation of the site (side chains only) in the Y137A mutant demonstrated by evident distortion of the Y48, Q133, and D135 conformations (Figure 5). Upon ligand binding, the original binding site topology needed to be reassumed and certain residues (*e.g.* D135) do not to fully reorient to their optimal positions.

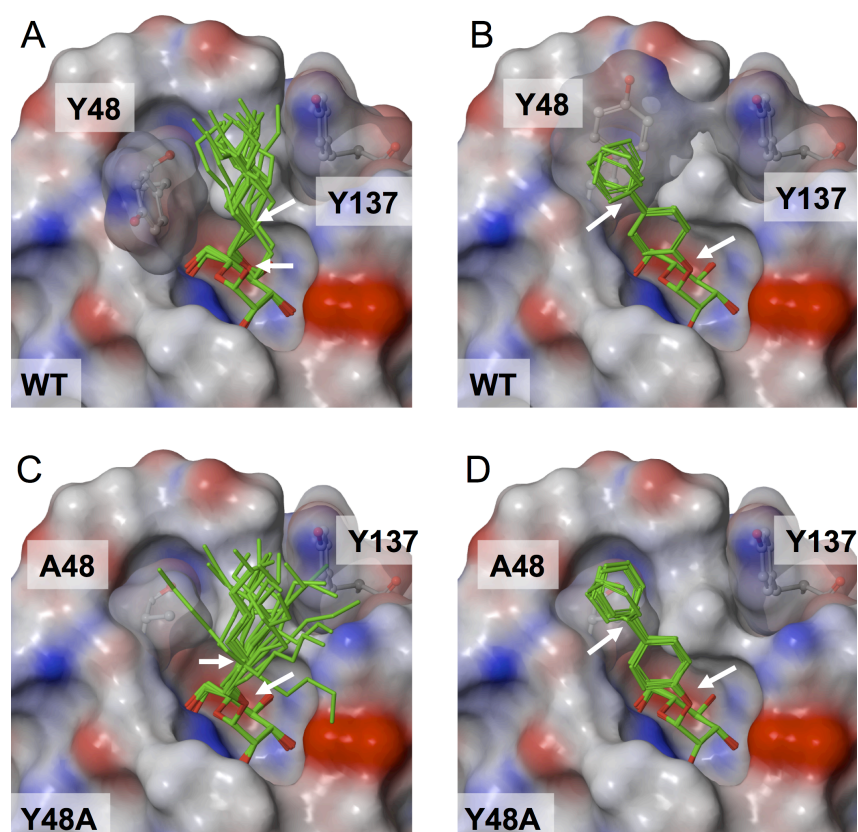


Figure 4: Molecular dynamic simulation of FimH-CRD wild type and the Y48A mutants: Surface representation of the binding site of the wild type (A) with *n*-heptyl α -D-mannopyranoside (2), and (B) with 4-biphenyl α -D-mannopyranoside (3) and Y48A (C) with *n*-heptyl α -D-mannopyranoside (2), and (D) with 4-biphenyl α -D-mannopyranoside (3) in their most commonly occurring conformations as indicated by simulation frame clustering. Constrained bounds are indicated by an arrow.

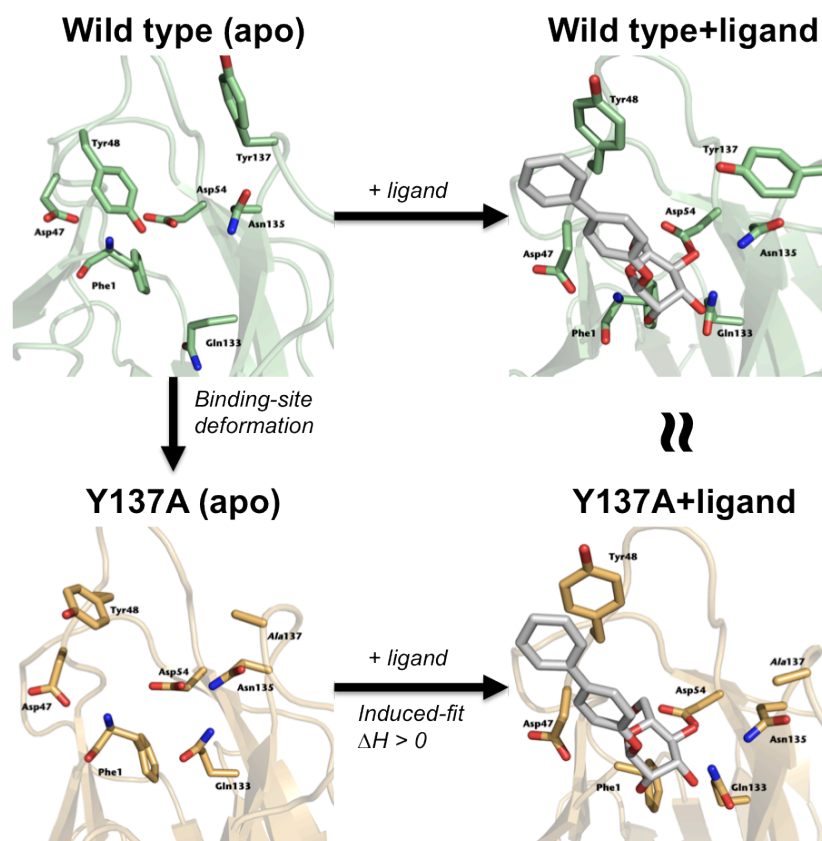


Figure 5: Snapshots from molecular dynamics simulations of FimH-CRD and FimH-CRD bound to the 4-biphenyl α -D-mannopyranoside (**3**) showing distortion of binding site residues in the Y137A mutant and the associated extra induced fit cost upon ligand binding.

Crystallographic study

In order to further confirm our findings, we obtained crystal structures of *n*-heptyl α -D-mannopyranoside in complex with the wild type, Y48A and Y137A FimH-CRD (Figure 6). Aside from Y48 side chain orientations impacted by crystal packing (as also reported for previous FimH-related studies in reference 31), the three structures supported our expectations regarding identical backbone and binding site conformations. The heptyl aglycones were found similarly oriented throughout all structures and interacting with both tyrosine side chains. For the Y48A mutant, alternative aglycone conformations could also be observed confirming increased ligand flexibility. Altogether the structures were found to closely resemble those obtained with molecular modeling and supported the corresponding predictions.

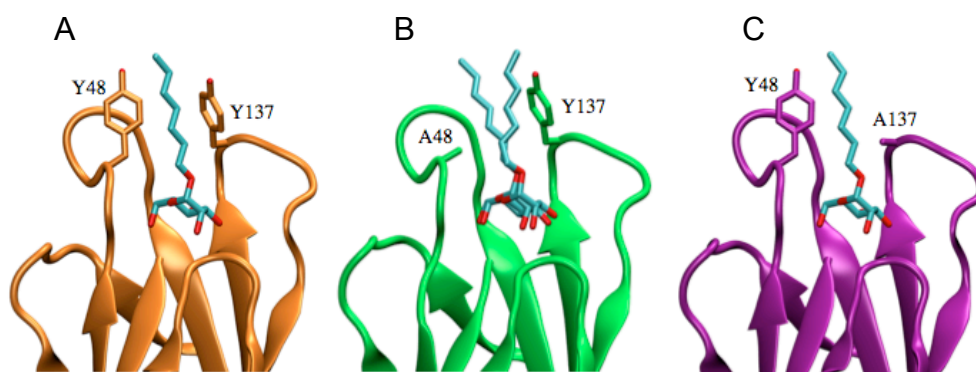


Figure 6: Co-crystal structures of FimH-CRD with *n*-heptyl α -D-mannopyranoside (**2**). FimH-CRD wild type (A), Y48A mutant (B) and Y137A mutant (C).

Discussion

Previous studies suggested that the affinity of FimH-CRD towards ligands such as alkyl and aryl mannosides was significantly improved due to favorable hydrophobic/ π - π interactions of the aglycone moiety with the phenyl rings of Y48 and Y137 residues (23,31). The *n*-butyl α -D-mannopyranoside was co-crystallized in an *in-docking* mode, where the butyl moiety interacts with both Y48 and Y137 residues (31). In contrast, for a meta-carboxyl ester substituted biphenyl α -D-mannopyranoside co-crystallized with the FimH-CRD, an *out-docking* mode was proposed (23). In this binding mode, π - π stacking of the outer aromatic ring of the biphenyl aglycone with Y48 was observed.

The objective of this study was to assess the contributions of the Y48 and Y137 residues towards the binding of various mannosides. We have therefore generated the alanine mutants Y48A and Y137A of FimH-CRD and confirmed that the mutations did not affect the secondary structure, and stability of the proteins. We obtained the corresponding affinities and thermodynamic binding profiles for the 1,5-anhydro D-mannitol (**1**), *n*-heptyl α -D-mannopyranoside (**2**) and the 4-biphenyl α -D-mannopyranoside (**3**). For the Y48A mutation, the rIC_{50} and K_D values for compound **2** and **3** were found only minimally affected compared to the wild type (Figure 2B & Table 1). Particularly, the result for compound **3** was not expected, given the generally accepted importance of the Y48 side chain towards alkyl- and aryl mannoside binding. Further inquiries into the thermodynamic fingerprints of these systems (obtained with ITC) showed reduced binding enthalpies compensated by enhanced entropy contributions (Table 1, Figure 3). Subsequent binding free energy calculations (Figure S3) revealed (as initially expected) that the loss of enthalpy originated from suppressed interactions with residue 48. The beneficial entropy changes were found (by means of molecular dynamics simulations) related to a lesser restriction of the binding interface unique to the non-natural mutation. Changes in the enthalpy can be interpreted in terms of differences in protein solvation and direct interactions between the ligand and protein. These interactions consist of hydrogen bonding-, electrostatic-, and van der Waals interactions, which can be identified based on computational methods and the structure of ligand-protein complex (58). The source of the entropy contribution, however, remains difficult to identify.

Based on available crystallographic data, the biphenyl aglycone of mannoside derivative **3** was predicted to interact primarily with the Y48 side chain. Consequently, the mutation Y137A was not foreseen to affect its affinity towards the FimH-CRD. Surprisingly, the competitive binding assay and thermodynamic data demonstrated that the Y137A substitution significantly diminished the affinity of compound **3**. The modeling studies showed significant deformation of the binding site of Y137A mutant as demonstrated by the distortion of Y48, Q133, and D135 side chains (Figure 5). However, this side chain re-orientation does not lead to a conformational change of the protein backbone. Upon ligand binding, the binding site conformation needs to be reassumed which incurs additional enthalpy

cost. Furthermore, certain residues (*e.g.* D135) appear unable to fully reorient, resulting in weaker interaction with the mannose moiety in the Y137A mutant. A comparable enthalpic loss was also observed in the case of *n*-heptyl α -D-mannopyranoside (**2**) and could be attributed to a similar induced fit penalty and/or to the loss of some direct interactions between the ligand and Y137 (data not shown).

Based on the presented data we conclude that both tyrosine gate residues are of crucial importance to the binding of alkyl and aryl mannosides to FimH, though their respective roles may vary between strongly interacting (Y48) and interacting/stabilizing the binding site conformation (Y137). In light of this, it would seem appropriate to treat FimH ligands on individual basis (and allowing the thought that they might maintain multiple conformations in the bound state) rather than classifying them between distinct, rigid binding modes. We believe that it is unlikely (given the corresponding entropic and desolvation penalties) that the flexible mannoside aglycones would favor rigid, directional interactions with more distant residues to those with the tyrosine gate. Contacts like this may nevertheless be observed under non-physiological, crystallographic conditions.

References

- (1) Hooton T.M., Stamm W.E. (1997) Diagnosis and treatment of uncomplicated urinary tract infection *Infect. Dis. Clin. N. Am.* **11**, 551–581.
- (2) Foxman B. (2002) Epidemiology of urinary tract infections: Incidence, morbidity, and economic costs. *Am. J. Med.* **113**, 5S–13S.
- (3) Ronald A. (2002) The etiology of urinary tract infection: Traditional and emerging pathogens. *Am. J. Med.* **8**, 14S–19S.
- (4) Klemm P. (1992) FimC, a chaperone-like periplasmic protein of *Escherichia coli* involved in biogenesis of type 1 fimbriae. *Res. Microbiol.* **143**, 831–838.
- (5) Martinez J.J., Mulvey M.A., Schilling J.D., Pinkner J.S., Hultgren S.J. (2000) Type 1 pilus-mediated bacterial invasion of bladder epithelial cells. *EMBO. J.* **19**, 2803–2812.
- (6) Wu X.R., Sun T.T., Medina J.J. (1996) *In vitro* binding of type 1-fimbriated *Escherichia coli* to uroplakins Ia and Ib: Relation to urinary tract infections. *Proc. Natl. Acad. Sci. USA.* **93**, 9630–9635.
- (7) Zhou G., Mo W.J., Sebbel P., Min G.W., Neubert T.A., Glockshuber R., Wu X.R., Sun T.T. (2001) Uroplakin Ia is the urothelial receptor for uropathogenic *Escherichia coli*: Evidence from *in vitro* FimH binding. *J. Cell Sci.* **114**, 4095–4103.
- (8) Xie B., Zhou G., Chan S.Y., Shapiro E., Kong X.P., Wu X.R., Sun T.T., Costello C.E. (2006) Distinct glycan structures of uroplakins Ia and Ib: Structural basis for the selective binding of FimH adhesin to uroplakin Ia. *J. Biol. Chem.* **281**, 14644–14653.
- (9) Anderson G.G., Palermo J.J., Schilling J.D., Roth R., Heuser J., Hultgren S.J. (2003) Intracellular bacterial biofilm-like pods in urinary tract infections. *Science* **301**, 105-107.
- (10) Mulvey M.A., Schilling J.D., Hultgren S.J. (2001) Establishment of a persistent *Escherichia coli* reservoir during the acute phase of a bladder infection. *Infect. Immun.* **69**, 4572-4579.
- (11) Mysorekar I.U., Hultgren S.J. (2006) Mechanisms of uropathogenic *Escherichia coli* persistence and eradication from the urinary tract. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 14170-14175.
- (12) Thomas W.E., Trintchina E., Forero M., Vogel V., Sokurenko E.V. (2002) Shear-dependent ‘stick-and-roll’ adhesion of type 1 fimbriated *Escherichia coli*. *Cell*, **109**, 913–923.
- (13) Aprikian P., Tchesnokova V., Kidd B., Yakovenko O., Yarov-Yarovoy V., Trinchina E., Vogel V., Thomas W., Sokurenko E. (2007) Interdomain interaction in the FimH adhesion of *Escherichia coli* regulates the affinity to mannose. *J. Biol. Chem.* **282**, 23437–23446.
- (14) Naber K.G. (2000) Treatment options for acute uncomplicated cystitis in adults. *J. Antimicrob. Chemother.* **46**, 23-27.
- (15) Warren J.W. (2001) Practice guidelines for the treatment of uncomplicated cystitis. *Curr. Urol. Rep.* **2**, 326-329.
- (16) Nicolle L.E. (2002) Urinary tract infection: traditional pharmacologic therapies. *Am. J. Med.* **113**, 35-44.

- (17) Stewart, P. S., and Costerton J. W. (2001) Antibiotic resistance of bacteria in biofilms. *Lancet* **358**, 135-138.
- (18) Soto, S. M., Smithson A., Horcajada J. P., Martinez J. A., Mensa J. P., Vila J. (2006) Implication of biofilm formation in the persistence of urinary tract infection caused by uropathogenic *Escherichia coli*. *Clin. Microbiol. Infect.* **12**, 1034-1036.
- (19) Foxman B. (2003) Epidemiology of urinary tract infections: incidence, morbidity, and economic costs. *Dis. Mon.* **49**, 53–70.
- (20) Firon N., Ashkenazi S., Mirelman D., Ofek I., Sharon N. (1987) Aromatic alpha-glycosides of mannose are powerful inhibitors of the adherence of type 1 fimbriated *Escherichia coli* to yeast and intestinal epithelial cells. *Infect. Immun.* **55**, 472-476.
- (21) Lindhorst T. K., Kötter S., Kubisch J., Krallmann-Wenzel U., Ehlers S., Kren V. (1998) Effect of *p*-substitution of aryl α -D-mannosides on inhibiting mannose-sensitive adhesion of *Escherichia coli*: Syntheses and testing. *Eur. J. Org. Chem.*, 1669-1674.
- (22) Sperling O., Fuchs A., Lindhorst T.K. (2006) Evaluation of the carbohydrate recognition domain of the bacterial adhesion FimH: design, synthesis and binding properties of mannoside ligands. *Org. Biomol. Chem.*, **4**, 3913-3922.
- (23) Han Z., Pinker J.S., Ford B., Obermann R., Nolan W., Wildman S.A., Hobbs D., Ellenberger T., Cusumano C.K., Hultgren S.J., Janetka, J. W. (2010) Structure-based drug design and optimization of mannoside bacterial FimH antagonists. *J. Med. Chem.* **53**, 4779-4792.
- (24) Klein T., Abgottspon D., Wittwer M., Rabbani S., Herold J., Jiang X., Kleeb S., Lüthi C., Scharenberg M., Bezencçon J., Gubler E., Pang L., Smiesko M., Cutting B., Schwardt O., Ernst B. (2010) FimH antagonists for the oral treatment of urinary tract infections: from design and synthesis to in vitro and in vivo evaluation. *J. Med. Chem.* **53**, 8627-8641.
- (25) Schwardt O., Rabbani S., Hartmann M., Abgottspon D., Wittwer M., Kleeb S., Zalewski A., Smieško M., Cutting B., Ernst. B. (2011) Design, synthesis and biological evaluation of mannosyl triazoles as FimH antagonists. *Bioorg. Med. Chem.* **19**, 6454-6473.
- (26) Cusumano C.K., Pinkner J. S., Han Z., Greene S.E., Ford B.A., Crowley J.R., Henderson J.P., Janetka J.W., Hultgren S.J. (2011) Treatment and prevention of urinary tract infection with orally active FimH inhibitors. *Sci. Transl. Med.* **3**, 109-115.
- (27) Pang L., Kleeb S., Lemme K., Rabbani S., Scharenberg M., Zalewski A., Schaedler F., Schwardt O., Ernst B. (2012) Fimh antagonists: Structure-activity and structure-property relationships for biphenyl α -D-mannopyranosides. *Chem. Med. Chem.* **7**, 1404-1422
- (28) Jiang X., Abgottspon D., Kleeb S., Rabbani S., Scharenberg M., Wittwer M., Haug M., Schwardt O., Ernst B. (2012) Antiadhesion therapy for urinary tract infections: A balanced PK/PD profile proved to be key for success. *J. Med. Chem.* **55**, 4700-4713

- (29) Choudhury D., Thompson A., Stojanoff V., Langermann S., Pinkner J., Hultgren S.J., Knight S.D., (1999) X-ray structure of the FimC-FimH chaperone-adhesin complex from uropathogenic *Escherichia coli*. *Science* **285**, 1061-1066.
- (30) Hung C.S., Bouckaert J., Hung D., Pinkner J., Widberg C., DeFusco A., Auguste C.G., Strouse R., Langermann S., Waksman G., Hultgren S.J. (2002) Structural basis of tropism of *Escherichia coli* to the bladder during urinary tract infection. *Mol. Microbiol.* **44**, 903–915.
- (31) Bouckaert J., Berglund J., Schembri M., Genst E.D., Cools L., Wuhner M., Hung C.S., Pinkner J., Slättergard R., Zavialov A., Choudhury D., Langermann S., Hultgren S.J., Wyns L., Klemm P., Oscarson S., Knight S.D., Greve H.D. (2005) Receptor binding studies disclose a novel class of high-affinity inhibitors of the *Escherichia coli* FimH adhesin. *Mol. Microbiol.* **55**, 441-455.
- (32) Wellens A., Garofalo C., Nguyen H., Van Gerven N., Slättergard R., Hernalsteens J.P., Wyns L., Oscarson S., De Greve H., Hultgren S., Bouckaert J. (2008) Intervening with urinary tract infections using anti-adhesives based on the crystal structure of the FimH-oligomannose-3 complex. *PLoS One* **3**, e2040.
- (33) Wellens A., Lahmann M., Touaibia M., Vaucher J. Oscarson S., Roy R., Remaut H., Bouckaert J. (2012) The tyrosine gate as a potential entropic lever in the receptor binding site of the bacterial adhesin FimH. *Biochem.* **51**, 4790-4799.
- (34) Ella-Menye J.R., Nie X., and Wang G. (2008) Synthesis of octahydropyrano(3,2-b)pyrrole-2-carboxylic acid derivatives from D-mannose. *Carbohydr. Res.* **343**, 1743–1753.
- (35) Oscarson S., Tiden A.K. (1993) Syntheses of the octyl and tetradecyl glycosides of 3,6-di-o-alpha-deuteriummannopyranosyl-alpha-deuterium-mannopyranose and of 3,4-di-o-alpha-deuterium-mannopyranosyl-alpha-deuterium-mannopyranose a new way for 2,4-di-o-protection of mannopyranosides. *Carbohydr. Res.*, **247**, 323–328.
- (36) Rabbani S., Jiang X., Schwardt O., Ernst B. (2010) Expression of the carbohydrate recognition domain of FimH and development of a competitive binding assay. *Anal. Biochem.* **407**, 188-195.
- (37) Ling M.M., and Robinson B.H. (1997) Approaches to DNA mutagenesis: an overview, *Anal. Biochem.* **254**, 157-178.
- (38) Meerman H.J., Georgiou G. (1994) Construction and characterization of a set of *Escherichia coli* strains deficient in all known loci affecting the proteolytic stability of secreted recombinant proteins. *Biotechnology* **12**, 1107–1110
- (39) Vetsch M., Sebbel P., Glockshuber R. (2002) Chaperone-independent folding of type 1 pilus domains. *J. Mol. Biol.* **322**, 827–840
- (40) Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
- (41) Bitsch F., Aichholz R., Kallen J., Geisse S., Fournier B., Schlaeppi J.M. (2003) Identification of natural ligands of retinoic acid receptor-related orphan receptor α ligand-binding domain expressed in Sf9 cells - a mass spectrometry approach. *Anal. Biochem.* **323**, 139-149.

- (42) Beychok S. (1966) Circular dichroism of biological macromolecules. *Science* **154**, 1288–1299.
- (43) Greenfield N.J. (2006) Using circular dichroism spectra to estimate protein secondary structure. *Nat. protocols* **1**, 2876–90.
- (44) Santoro M.M., Bolen D.W. (1988) Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants. *Biochemistry*, **27**, 8063–8068.
- (45) Pace C.N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves. *Methods Enzymol.* **131**, 266–280.
- (46) Ziegler A. and Seelig J. (2004) Interaction of the protein transduction domain of HIV-1 TAT with heparan sulfate: Binding mechanism and thermodynamic parameters. *Biophys. J.* **86**, 254–263.
- (47) Wiseman T., Williston S., Brandts J. F., Lin L.N. (1989) Rapid measurement of binding constants and heats of binding using a new titration calorimeter. *Anal. Biochem.* **179**, 131–137.
- (48) Bowers K.J., Chow E., Xu H., Dror R.O., Eastwood M.P., Gregersen B.A., Klepeis J.L., Kolossvary I., Moraes M.A., Sacerdoti F.D., Salmon J.K., Shan Y., and Shaw D.E. (2006) Proceedings of the ACM/IEEE conference on supercomputing (SC06), Tampa, FL, November 11–17, 2006, ACM, New York.
- (49) Jorgensen W.L., Chandrasekhar J., Madura J.D., Impey R.W., and Klein M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
- (50) Darden T., York D., and Pedersen L. (1993) Particle mesh Ewald: An $N \log(N)$ method for Ewald sums in large systems. *J. Chem. Phys.* **98**, 10089–10092.
- (51) Ryckaert J.P., Ciccotti G., and Berendsen J.C. (1977) Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *J. Comput. Phys.* **23**, 327–341
- (52) Martyna G.J., Tobias D.J., and Klein, M.L. (1994) Constant pressure molecular dynamics algorithms. *J. Chem. Phys.* **101**, 4177– 4189.
- (53) Nosé S. (1984) A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.* **52**, 255–268.
- (54) DeLano W. L. (2010) The PyMOL Molecular Graphics System, Version 1.3r1, Schrodinger, LLC, New York.
- (55) Schüttelkopf A.W. and van Aalten D.M.F. (2004) PRODRG: a tool for high-throughput crystallography of protein-ligand complexes. *Acta Crystallogr. Sect. D.* **60**, 1355–1363.
- (56) Adams P. D., Grosse-Kunstleve R. W., Hung L. W., Ioerger T. R., McCoy A. J., Moriarty N. W., Read R. J., Sacchettini J. C., Sauter N. K., Terwilliger T. C. (2002) PHENIX: building new software for automated crystallographic structure determination. *Acta Crystallogr. Sect. D.* **58**, 1948–1954.
- (57) Durka M., Buffet K., Iehl J., Holler M., Nierengarten J.F., Taganna J., Bouckaert J., and Vincent S.P. (2011) The functional valency of dodecamannosylated fullerenes with *Escherichia coli* FimH--towards novel bacterial antiadhesives. *Chem. Commun.* **47**, 1321–1323.

(58) Imberty A., Hardman K.D., Carver J.P., and Perez S. (1991) Molecular modeling of protein-carbohydrate interactions. Docking of monosaccharides in the binding site of concanavalin A. *Glycobiology* **16**, 631-642.

Supplementary Material

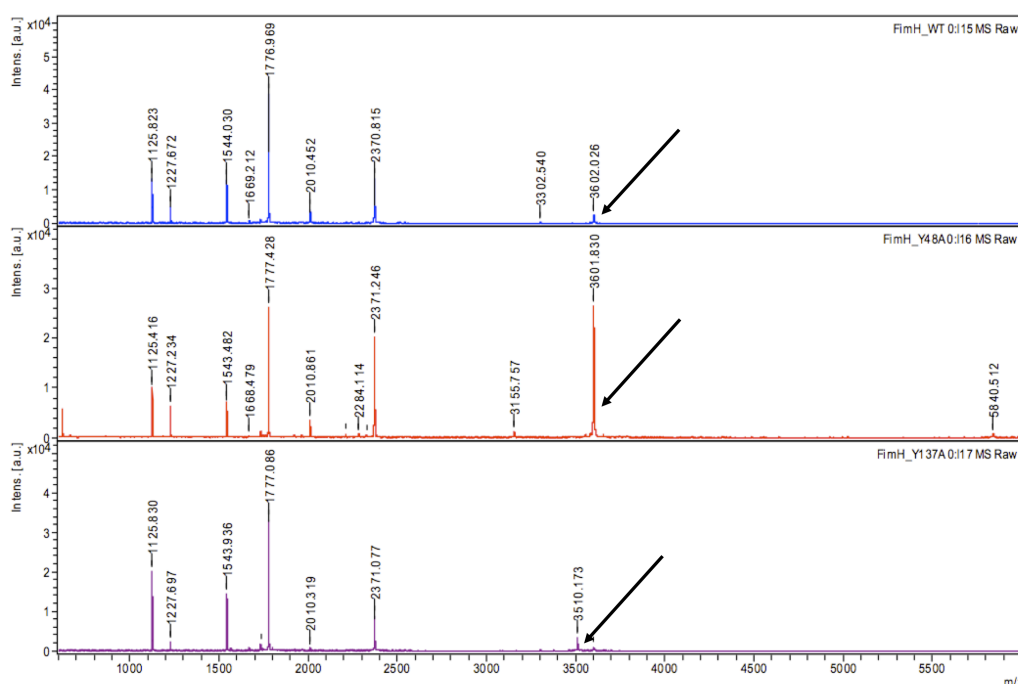


Figure S1. To differentiate between Y48A and Y137A mutants the proteins were subjected to trypsin digestion followed by MS analysis. The arrows indicate the 31 amino acid peptide with the same molecular weight in the wild type and Y48A mutants, and with a molecular weight 92 Da smaller in the Y137A mutant.

MM-GBSA post-processing—To gain insight into binding thermodynamics of respective protein-ligand systems, ensembles of equally spaced (120 ps) snapshots were subjected to the MM-GBSA free energy calculation (Prime, version 3.0, Schrödinger, LLC, New York, NY, 2011.). The energies were computed as mean values using equation S1.

$$\Delta G_{bind} = G_{complex} - (G_{receptor} + G_{ligand}) \text{ (Eq. S1)}$$

Structures of each molecular species were obtained by decomposing corresponding snapshots. All calculations were performed using a continuum solvent model. Because the method does not account for large, unfavorable changes in solute conformational entropies, the obtained results do not scale with the experimental data and reflect mostly the enthalpic changes for each system.

Table S1: ΔG values [kcal/mol] calculated over ensembles of 250 MD snapshots along with corresponding standard errors of the means (SEM).

	WT		Y48A		Y137A	
	ΔG	SEM	ΔG	SEM	ΔG	SEM
1,5 Anhydro D-mannitol (1)	−48.2	0.3	−48.4	0.2	−50.0	0.3
<i>n</i> -Heptyl α -D-mannopyranoside (2)	−65.9	0.4	−60.7	0.4	−60.8	0.4
4-Biphenyl α -D-mannopyranoside (3)	−65.6	0.4	−61.1	0.4	−58.0	0.3

Analysis of FimH MD simulations—We developed an in-house script to post-process MD simulations of FimH (free and in complex with different ligands) and highlight statistically significant differences between them. The python script computes various geometrical and energetic parameters for snapshots along the MD simulation using Schrödinger Suite API (Schrödinger, LLC, New York, NY, 2011). Afterwards, statistical analysis was performed on the collected time series data and a difference map was generated for every studied pair of simulations. Table S1 lists studied parameters and explains their calculation methodologies and intended purposes.

Difference map for free FimH wild type vs. Y137A mutant—Table S2 lists parameters that showed statistically significant variation upon the Y137A mutation in FimH binding-site. In our study, binding-site residues were defined as residues with at least one atom within 8 Å distance from any ligand atom. For the ligand free simulations binding-site residues were defined based on the position of 4-biphenyl α -D-mannopyranoside (**3**) (out mode). Our assessment of the effects of Y137A mutation indicated an important role for the Y137 residue in stabilizing binding-site conformation. When Y137 was mutated into alanine, the FimH binding-site suffered from obvious distortion reflected by noticeable changes in protein–2d intramolecular binding-site interaction energies as well as several binding-site COM distances. Additionally, several dihedral angles (backbone and side chain) demonstrated statistically significant alteration between the wild type and the Y137A mutant. Results indicate that Y48 is one of the most affected residues by the studied mutation. The observed distortion also involves other binding-site residues such as Phe1, Ile52, Ile13, and Asn138. However, it is difficult to link the observed binding-site distortion to specific backbone and/or side chain dihedral angle(s).

Table S2. Geometrical and energetic parameters used for assessment of differences between MD simulations.

Parameter name	Computation methodology	Purpose
phi angle	Geometrical	Assessment of backbone rearrangement
psi angle	Geometrical	Assessment of backbone rearrangement
chi angle	Geometrical	Assessment of side chain conformational distribution
COM distance	Geometrical, distance (in Å) between centers-of-mass of pairs of binding-site residues	Assessment of overall conservation of binding-site conformation
ligand–protein	Sum of non-bonded forcefield terms (OPLS_2005) between ligand and individual binding-site residues (MacroModel, version 9.9, Schrödinger, LLC, New York, NY, 2011)	Assessment of loss or gain of interaction(s) between the ligand and binding-site residue(s)
protein–2d	Intramolecular interaction energies between pairs of binding-site residues calculated using the same approach used for ligand–protein parameter	Assessment of conservation of proper binding-site conformation. A collective parameter for phi/psi/chi angles and COM distance parameters, since any changes in binding-site geometry, on the level of one or more residues, would be reflected in a change in its interaction energy with neighboring residue(s).

Table S3. Difference map for ligand free FimH wild type vs. Y137A mutant highlighting statistically significant changes in binding-site residues.

Parameter	Average (wild type)	Average (Y137A mutant)	Difference (S.D.)
phi angles (degrees)			
Gly15	205.8	113.9	−92.0 (1.4)
Ala10	275.1	243.8	−31.3 (0.7)
Gly16	158.9	134.1	−24.7 (1.3)
Asp47	242.6	261.6	19.0 (0.5)
Asn136	269.4	287.1	17.7 (0.5)
Gly73	219.6	195.5	−24.0 (1.4)
Val75	257.2	234.6	−22.7 (0.5)
Ala18	233.3	215.0	−18.3 (0.7)
Asn138	221.7	239.0	17.3 (0.7)
Ile13	284.6	300.3	15.7 (0.6)
Lys101	245.3	259.9	14.6 (0.6)
Ile52	231.8	243.9	12.2 (0.6)
psi angles (degrees)			
Gly16	251.7	202.7	−48.9 (1.5)
Val30	116.3	90.1	−26.1 (1.2)
Asp47	9.2	346.4	−22.9 (0.5)
Gly15	172.5	193.3	20.8 (1.4)
Pro12	150.3	170.4	20.1 (0.8)
Ile52	136.2	117.4	−18.8 (0.5)
Ile42	103.5	121.6	18.0 (0.7)
Ser72	140.4	158.0	17.6 (0.9)
Asn70	326.4	342.3	15.9 (0.6)
Phe144	127.3	111.5	−15.7 (0.5)
chi angles (degrees)			
Tyr55-2	292.3	115.5	−176.8 (2.1)
Asp54-2	301.2	114.8	173.5 (1.5)
Asn138-1	58.7	248.9	−169.7 (1.0)
Thr99-1	189.7	55.6	−134.1 (0.5)
Asp141-2	272.8	167.9	−104.9 (2.4)
Tyr48-1	46.3	302.4	−103.9 (0.8)
Asp140-1	36.6	294.0	−102.6 (1.3)
Phe142-1	205.9	303.5	97.6 (1.5)
Thr53-1	298.7	34.6	95.9 (1.0)
Ile11-2	250.1	176.6	−73.5 (0.5)
Ile52-2	167.7	231.2	63.5 (1.4)
Tyr48-2	41.9	97.9	56.0 (1.5)
Thr51-1	103.8	56.4	−47.4 (0.5)
protein-2d energies (kJ/mol)			
His45-Asp47	−15.7	−24.2	−8.4 (0.5)

Phe1-Tyr48	-10.5	-2.8	7.7 (0.2)
Thr53-Asn136	-19.1	-11.3	7.8 (0.4)
Asn46-Tyr48	-23.9	-18.9	5.1 (0.2)
Phe1-Gln133	-12.4	-17.3	-4.9 (0.2)
Tyr48-Ile52	-21.9	-17.1	4.7 (0.2)
Asn138-Phe142	-0.7	-4.8	-4.1 (0.1)
Phe1- Phe144	-4.0	-8.0	-4.0 (0.1)
Ile13-Phe144	-7.2	-3.2	3.9 (0.2)
Thr53-Tyr55	-15.2	-11.7	3.6 (0.2)
Leu131-Phe144	-27.2	-30.6	-3.4 (0.2)
Ile13-Phe142	-5.1	-1.7	3.4 (0.2)
Asn46-Thr53	-1.3	-4.1	-2.8 (0.1)
Tyr48-Asn135	-3.3	-0.5	2.8 (0.1)
Asn46-Ile52	-9.2	-11.9	-2.7 (0.2)
Thr53-Asn135	-11.4	-9.2	2.3 (0.1)
COM distances (Å)			
Asn138-Phe142	10.1	7.8	-2.3 (0.04)
Tyr48-Arg132	16.4	18.5	2.2 (0.04)
Tyr48-Tyr55	11.2	13.4	2.2 (0.04)
Tyr48-Thr134	11.6	13.7	2.1 (0.04)
Tyr48-Val56	13.5	15.7	2.2 (0.03)
Ile13-Asn138	14.7	12.7	-2.0 (0.06)
Tyr48-Leu131	16.5	18.3	1.7 (0.03)
Tyr48-Tyr95	10.8	12.5	1.7 (0.03)
Tyr48-Thr53	7.3	9.1	1.7 (0.04)
Tyr48-Asn135	8.6	10.3	1.7 (0.04)
Tyr48-Asn136	9.5	11.1	1.6 (0.03)
Tyr48-Gln133	10.9	12.5	1.6 (0.04)
Asn136-Phe142	14.2	12.6	-1.6 (0.03)
Asn138-Phe144	16.1	14.5	-1.6 (0.03)
Asn135-Phe142	9.7	8.1	-1.6 (0.03)
Ile13-Thr51	17.9	16.5	-1.4 (0.06)
Ile13-Ile52	15.1	13.7	-1.4 (0.06)
Phe1-Tyr48	8.2	9.5	1.4 (0.02)

Difference map for 4-biphenyl α -D-mannopyranoside (3) in complex with FimH-CRD wild type vs. Y137A mutant—For the purpose of this study, MD simulations were carried out for the wild type FimH in complex with 4-biphenyl α -D-mannopyranoside (**3**) adopting the *out* binding mode observed in crystal structure of an analogous derivative (PDB: 3MCY). The Y137A mutant was created starting from the same pose and relaxed using the default relaxation protocol in Desmond MD package. Table S3 lists parameters that showed statistically significant variation upon the Y137A mutation in FimH complex with antagonist **3**. The resulting difference map shows a relatively smaller number of differences compared to the ligand free simulations.

The comparatively limited differences associated with Y137A mutation when antagonist **3** is bound indicate that in presence of the ligand the binding-site *maintains* its conformation relative to the wild type. For instance, a significant distortion of Y48 at the tyrosine gate was observed upon mutating Y137 to alanine in the ligand free simulations. Notably, no sign of this distortion was observed in antagonist **3**-bound simulations. This led us to the conclusion that the *bioactive* binding-site conformation is more or less maintained when a ligand is bound and that the distortion observed in the *apo* simulations needs to be first rectified to accommodate the ligand, which would incur an extra induced fit penalty as discussed in text. Difference map for ligand–protein interactions also revealed drop in interaction energies with Asn135, Asp47, and Gln133. The proposed readjustment of the Y137A mutant binding-site upon ligand binding, thus, seems to be incomplete.

Table S4. Difference map for the 4-biphenyl α -D-mannopyranoside bound to FimH-CRD wild type vs. Y137A mutant highlighting statistically significant changes in binding-site residues.

Parameter	Average (wild type)	Average (Y137A mutant)	Difference (S.D.)
phi angles (degrees)			
Asn138	273.4	216.4	−57.0 (0.6)
Asp140	227.4	275.5	48.1 (0.8)
Ser139	238.6	269.5	31.0 (0.6)
Asp141	270.6	241.9	−28.7 (0.7)
Asn136	295.4	270.3	−25.1 (0.5)
psi angles (degrees)			
Asp140	125.8	151.4	25.6 (0.6)
Gly61	100.2	118.3	18.2 (0.6)
Phe144	104.9	120.1	15.2 (0.4)
chi angles (degrees)			
Asn138-1	185.2	58.0	−127.2 (0.6)
Asp141-2	47.3	323.0	−84.3 (2.4)
Phe142-1	302.7	222.2	−80.5 (1.0)
Glu50-2	77.5	43.8	−33.7 (1.4)
Asp140-1	88.5	35.3	−53.2 (1.9)
protein-2d energies (kJ/mol)			
Phe1-Asp47	−120.9	−112.2	8.8 (0.4)
His45-Asp47	−25.7	−18.8	6.9 (0.5)
Asp141-Phe142	−142.4	−136.4	6.1 (0.2)
Gln133-Phe142	−32.5	−27.6	4.8 (0.2)
Asn135-Asn136	−128.3	−123.5	4.8 (0.3)
Asn138-Phe142	−4.3	−0.5	3.8 (0.1)
Phe1-Phe144	−10.0	−6.4	3.6 (0.1)
Asp54-Asn135	−17.8	−14.7	3.0 (0.1)
Asp54-Gln133	−32.8	−38.7	−5.8 (0.4)
ligand–protein energies (kJ/mol)			
lig-Phe1	−121.5	−108.8	12.7 (0.5)
lig-Ile13	−10.0	−1.3	8.7 (0.1)
lig-Asn135	−19.0	−12.0	6.9 (0.3)
lig-Gln133	−15.2	−10.6	4.6 (0.2)
lig-Asp47	−14.7	−18.0	−3.3 (0.2)
lig-Phe142	−4.1	−1.4	2.6 (0.1)
COM distances (Å)			
Asn138-Phe142	7.7	10.3	2.5 (0.03)
Asn136-Phe142	12.6	14.4	1.8 (0.03)
Ser139-Phe142	9.3	11.1	1.8 (0.03)
Asn135-Asp140	8.8	7.1	−1.7 (0.03)
Asn135-Phe142	8.2	10.0	1.7 (0.03)
Thr134-Phe142	6.7	8.3	1.6 (0.02)
Thr51-Phe142	16.8	18.3	1.5 (0.03)
Asp47-Asn138	15.0	13.5	−1.5 (0.02)

Table S5. Data collection and refinement statistics

	FimH wild type with <i>n</i>-Heptyl α-D-mannopyranoside	FimH Y48A with <i>n</i>-Heptyl α-D-mannopyranoside	FimH Y137A with <i>n</i>-Heptyl α-D-mannopyranoside
Crystallization conditions	5% PGA-LM 0.1M Tris-HCl pH 7.8 30% v/v PEG 550 MME	5% w/v PGA-LM 100 mM Na cacodylate pH 6.5 200 mM MgCl ₂	5% PGA-LM 100 mM Na cacodylate pH 6.5 12 % w/v PEG 8K
Ligand	10 mM <i>n</i> -Heptyl α -D-mannopyranoside	10 mM <i>n</i> -Heptyl α -D-mannopyranoside	10 mM <i>n</i> -Heptyl α -D-mannopyranoside
Data collection			
X-ray wavelength (Å)	0.918410	0.96	0.918410
Beamline/Synchrotron	PX14.2, BESSY II	Proxima 1, SOLEIL	PX14.2, BESSY II
Resolution range (Å)	45 – 2.20 Å	41 - 3.01 Å	15 - 1.40 Å
Number of reflections (of which unique)	146524 (20238)	65358 (9163)	170620 (25691)
<I/sigI>	12.33 (3.77)	8.58 (3.02)	17.03 (4.15)
Completeness (%)	98.0 (90.8)	97.7 (99.1)	98.4 (91.8)
Crystal mosaicity (°)	0.448	0.111	0.445
R _{meas} ^b (%)	13.3 (57.1)	22.0 (81.1)	9.1 (47.4)
Wilson B-factor	35.92 (sfcheck)	43.085 (correct)	9.50 (scala)
Space group	P 2 ₁ 2 ₁ 2 ₁	P 4 ₃ 2 ₁ 2	P 2 2 ₁ 2 ₁
# molecules / a.u.	2	2	1
Unit cell dimensions			
a (Å)	60.305	89.850	45.974
b (Å)	68.070	89.850	59.552
c (Å)	95.607	91.870	96.885
α, β, γ (°)			- 1.70
Refinement			
R _{work}	0.2042	0.1787	0.1493
R _{free} (test set = 5%)	0.2501	0.2346	0.1829
R.m.s.d. stereochemistry			
Bonds (Å)	0.013	0.015	0.006
Angles (°)	1.474	1.712	1.232
Residues in favored and allowed regions (%) of Ramachandran	96.83 3.17	96.90 6.18	96.20 3.80
PDB entry	4BUQ	4CA4	Unpublished

Chapter 3.5. – Manuscript 7

Naturally Occurring Variations of the FimH Adhesin: Impact on Antagonist Binding

Formatted according to the requirements of *Journal of Biological Chemistry*.

Contributions

- | | |
|----------------|---|
| R. C. Preston | <ul style="list-style-type: none">• Generation of isogenic strains expressing different FimH variants.• Development and performance of static binding and inhibition assays.• Establishment and performance of flow assays.• Manuscript preparation. |
| M. Scharenberg | <ul style="list-style-type: none">• Sequencing of clinical isolates.• Type 1 pili immunostaining.• Electron microscopy. |
| D. Eriş | <ul style="list-style-type: none">• Generation of isogenic strains expressing different FimH variants.• Type 1 pili immunostaining.• Electron microscopy. |
| F. Hulliger | <ul style="list-style-type: none">• Evaluation of ligand affinity and performance of flow assays. |
| D. Abgottspon | <ul style="list-style-type: none">• Sequencing of clinical isolates. |

Naturally Occurring Variations of the FimH Adhesin: Impact on Antagonist Binding

Roland C. Preston^{1**}, Meike Scharenberg^{1**}, Deniz Eris¹, Fabian Hulliger¹, Daniela Abgottspon¹, Beat Ernst^{1*}

¹ Institute of Molecular Pharmacy, Pharmacenter, Klingelbergstrasse 50/70, 4054 Basel, Switzerland

* Correspondence to Beat Ernst. Tel.: +41 61 257 15 50, E-mail: beat.ernst@unibas.ch

** R.C.P. and M.E. contributed equally to this work.

Keywords

Urinary tract infection; type I pili; FimH; antibacterial treatment; FimH antagonist.

Abbreviations

1M, monomannose; CFU, colony-forming units; LD, lectin domain; LIBS, ligand-induced binding site; RNase B, ribonuclease B; UPEC, uropathogenic *Escherichia coli*; UPIa, uroplakin Ia; UTI, urinary tract infection.

Acknowledgements

The authors acknowledge the financial support of the Swiss National Science Foundation (200020_129935/1). We would also like to thank the University Hospital Basel, Switzerland, for providing clinical UPEC isolates.

Capsule

Background: Effect of FimH variants as found in clinical isolates of uropathogenic *E. coli* on α -D-mannopyranoside based antagonist binding.

Results: All investigated antagonist classes were efficacious in preventing *E. coli* binding, however, with distinct behavior of a squaric acid substituted α -D-mannopyranoside.

Conclusion: FimH antagonists might be successful therapeutics for uropathogenic *E. coli* with different naturally occurring mutations in FimH.

Significance: This is the first study to date that demonstrates efficacy of different FimH antagonist classes to various types of FimH.

Summary

FimH is a bacterial lectin located on type 1 pili of uropathogenic *E. coli* (UPEC) that mediates shear-enhanced adhesion to mannosylated surfaces. Binding of UPEC to uroepithelial cells initiates urinary tract infections (UTI), one of the most prevalent infections worldwide. Anti-adhesive glycomimetics based on α -D-mannopyranosides present an effective alternative to antibiotics for UTI treatment. However, genetic variation of *fimH* found in clinical isolated UPECs leads to distinct binding modes. FimH antagonists have not been systematically evaluated with these naturally occurring FimH variants to date.

In this study we investigated the genetic variation of *fimH* in 57 clinical UPEC isolates and demonstrated the distinct binding characteristics of four FimH variants to mannose-based ligands under static and hydrodynamic conditions. Furthermore, we evaluated binding potencies of four FimH antagonist representing four different classes.

We showed that the FimH variants exhibited individual binding behavior, thus covering a range of FimH binding phenotypes (low to strong binding) found in patients with cystitis. Inhibition assays showed that the FimH variants were sensitive to all antagonist classes. The alkyl-, indolinyphenyl-, and biphenyl derivatives showed a phenotype-dependent binding affinity, while the squaric acid derivative inhibited the different FimH variants independent of their binding phenotype. Our results demonstrate the efficacy of FimH antagonists towards the clinically relevant FimH variants.

Introduction

Urinary tract infections are one of the most prevalent bacterial infections with millions of people affected each year causing high costs in health care. In 70–95 % of all cases, UTIs are caused by uropathogenic *Escherichia coli* (UPEC) (1). The infection is initiated by the adhesion of UPECs through type 1 pili that are presented on the bacterial surface (2). These filamentous organelles consist of a rigid helical rod with up to 5000 of FimA subunits and a short tip with the subunits FimF, FimG, and FimH. It is the bacterial lectin FimH that confers the adhesion to bladder epithelial cells by allowing the interaction to mannosylated surface glycoproteins, mainly uroplakin Ia (UPIa) (3,4). FimH is composed of two domains, the pilin domain (FimH_{PD}) and lectin domain (FimH_{LD}) (5). The mannose binding pocket is located in the FimH_{LD}, while being anchored to the pili via the FimH_{PD}. Both domains have a beta-barrel fold. FimH has been shown to adopt two conformational states with distinct binding characteristics (5,6). In the unstrained low binding affinity conformation, interdomain bonds between the FimH_{PD} and FimH_{LD} compress the latter. The FimH_{LD} consequently displays a shallow binding groove that allows only for weak mannose binding. In the strained high binding affinity conformation, the loss of interdomain bonds induces a finger-trap like constriction of beta-sheets and an overall elongation of the FimH_{LD}. In this state, a deep mannose binding pocket is formed with a high affinity towards mannose. The switching between the two conformational states allows FimH to exhibit catch-bond behavior, *i.e.* the lifetime of a FimH-mannose bond is increased upon tensile force (7,8), an uncommon feature among protein-ligand bonds. Both the complexation of the chaperone FimC to FimH, as observed in FimC-FimH crystal structures (5), as well as the absence of the PD (9) allow the FimH_{LD} to adopt the high-affinity state due to the absence of the interdomain bonds. Using an antibody specific for the high-affinity state, Tchesnokova *et al.* demonstrated the allosteric connection between the high-affinity state and the open conformation of the interdomain region (10). Genetic variation in the *fimH* gene among UPEC strains was shown to alter mannose-binding affinities, despite the fact that the majority of all naturally occurring variations are not located in or in close proximity to the highly conserved mannose-binding site, but in the PD-LD interdomain region and (7,8,11,12). UPEC bearing the high-affinity form of FimH evolved under positive selection due to their enhanced virulence (11). At the same time, however, these UPEC strains are more sensitive to soluble inhibitors, *e.g.* the Tamm-Horsfall glycoprotein, and have a diminished ability to spread (13-15). Therefore, the catch bond behavior of FimH is believed to present a mechanism for UPEC to withstand

clearance from the bladder under tensile forces as they arise during micturition, while being insensitive towards soluble inhibitors (16).

Since the FimH-mediated adhesion of UPECs to the urothelium is the initial step of an infection, this bacterial lectin has drawn a lot of attention as drugable target for the treatment of UTIs. FimH antagonists provide a novel therapeutic opportunity for the treatment and prevention of UTIs, as opposed to traditional treatment with antibiotics which trigger bacterial resistance (17-19). Efforts in identifying potent FimH antagonists has lead to the identification of several classes of substituted α -D-mannopyranoside based compounds, e.g. alkyl- (9,20,21), biphenyl- (22,23), indolylphenyl- and indolinyphenyl- (24), and squaric acid (23) substituted α -D-mannopyranosides. The *in vitro* potency of FimH antagonists is usually determined in target-based assays using recombinantly expressed protein, either the FimH_{LD} alone (25) or full FimH in presence of FimC (26). In *in vitro* cell-based hemagglutination assays, a single clinical isolate (UTI89) with preferably high mannose affinity is commonly employed (27,28). Since UPEC strains cover a wide range of binding affinities towards mannose based upon their genetic variation, FimH antagonists are required to be efficacious to all binding states, a feature uninvestigated to date. We therefore identified commonly found mutations in UPECs isolated from patients with acute cystitis and generated corresponding isogenic *E. coli* strains expressing the FimH variants found therein. We further evaluated four classes of FimH antagonists for their binding affinity.

Experimental Procedures

Sequencing of clinical isolates—Clinical isolates were collected at the University Hospital Basel, Switzerland during 2009 and were selectively cultivated. Age and sex of the patient and the antibiotic resistances are known. The *fimH* gene was amplified by standard colony PCR using the two sequencing primers shown in Tab. 1. DNA samples were submitted for sequencing (Microsynth, Balgach, Switzerland). A list of all sequenced isolates and their antibiotic resistances are listed in Tab. S1.

TABLE 1.
Primers used in this study (5' to 3'). Mutations are underlined.

Sequencing Primers	
FimH seq forward	ACCGCGCAAAACATCCAGTT
FimH seq reverse	CCGGTGGCGCTTTATTTG
Flanking Primers	
FimH forward	GCGTGCACTCAGGGGAACCAT
FimH reverse	GCGCATGCTTATTGATAAACAAAAGT
V27A mutation	
V27A forward	CCCG <u>C</u> CGTGAATGTGGGGC
V27A reverse	CACGG <u>C</u> GGGCGCAAGGTTTAC
S62A mutation	
S62A forward	CAACGAGGCG <u>C</u> CGCTTATGGC
S62A reverse	GCCATAAGCG <u>G</u> CGCCTCGTTG
N70S_S78N mutation	
N70S_S78N forward	TTATCTAG <u>T</u> TTTTCCGGGACCGTAAATATAATGGCAGTAGC
N70S_S78N reverse	GCTACTGCCA <u>T</u> TATATTTTACGGTCCCGGAAAAA <u>C</u> TAGATAA
V128M mutation	
V128M forward	TTAATTGCCATGCTTATTTTGCGA
V128M reverse	TCGCAAAATAAGCAT <u>G</u> GGCAATTAA
A188D mutation	
A188D forward	ACCGTTTATTGTGAT <u>A</u> AAAAGCCAAAAC
A188D reverse	GTTTTGGCTTTTAT <u>C</u> ACAATAAACGGT

Generation of isogenic strains—The strains used herein are summarized in Tab. 2. The KB18 strain containing the pPKL114 plasmid and the ELT115 strain were kindly provided by Prof. Evgeni Sokurenko, University of Washington, Seattle, WA (29). The pPKL114 plasmid encodes the whole *fim* operon with a stop linker upstream of the *fimH* gene. The plasmid pGB2-24 containing *fimH*^{J96} was isolated from the ELT115 strain and served as template. Single nucleotide point mutations in FimH were introduced with the primers shown in Tab. 1 using overlap PCR following standard molecular techniques. The mutated *fimH* gene was inserted into *Apa*LI and *Sph*I (New England BioLabs, Allschwil, Switzerland) treated pGB2-

24 plasmid. The plasmid was amplified in the DH5 α *E.coli* strain (Novagen, Lucerne, Switzerland). The presence of mutations was confirmed by standard DNA sequencing (Microsynth, Balgach, Switzerland). The plasmids were transformed into the KB18 strain by electroporation. Successful transformation was verified by ampicillin (pPKL114) and/or chloramphenicol (pGB2-24) resistance.

TABLE 2
List of *E. coli* strains employed in this study.

FimH Variant	Background	Plasmids	Mutations	Original Name	Original Reference
Δ FimH	AAEC191A	pPKL114	-	KB18	(14)
FimH ^{J96}	KB18	pPKL114 pGB2-24	Wild Type	ELT115	(14)
FimH ^{F18}	KB18	pPKL114 pGB2-24	V27A, N70S, S78N	KB18	(14)
FimH ^{UT189}	KB18	pPKL114 pGB2-24	V27A, S62A, N70S, S78N	-	(30)
FimH ^{B126}	KB18	pPKL114 pGB2-24	V27A, V128M	-	(10)
FimH ^{high}	KB18	pPKL114 pGB2-24	A188D	-	(15)

Preparation of bacterial suspensions—The bacteria were grown overnight under shaking conditions in LB medium at 37 °C with appropriate antibiotics. Bacteria were washed three times with phosphate-buffered saline (PBS, pH 7.4) and diluted to final concentrations of 1·10⁹ CFU/ml for static assays, 5·10⁸ CFU/ml for competitive assays, and 1.5·10⁸ CFU/ml for flow chamber assays with PBS containing 0.2 % bovine serum albumin (BSA).

Type 1 pili immunostaining—Immunostaining was performed to verify equal expression of type 1 pili among the recombinant strains. Bacteria were incubated for 30 min at 25 °C with rabbit serum raised against type 1 pili (1:300 in PBS, 2 % BSA, kind gift from Prof. Dr. Krogfelt, Statens Serum Institut, Copenhagen, Denmark). After a washing step, the bacteria were treated with the secondary anti-rabbit Alexa647 conjugated antibody (10 µg/ml, Invitrogen, Lucerne, Switzerland) for 30 min at 25 °C. The bacteria were washed and resuspended in 300 µl PBS. Samples were measured with a CyAn ADP flow cytometer (BeckmanCoulter, Brea, CA). A total of 5·10⁴ bacteria were measured per sample. Data were acquired in a linear mode for the side scatter (SSC), a logarithmic mode for forward scatter (FSC) and the red fluorescent channel FL8-H (Alexa647). Quantification of fluorescent labeling was evaluated with FlowJo 7.6.3 software (Tree Star, Inc., Ashland, USA). The

KB18 strain was used as negative control, since it expresses only a few non-adhesive pili (31).

Bacterial binding under flow—Parallel plate flow chamber assays were performed similar as described previously (32). Cell culture dishes (35 mm, Corning, Inc., Corning, NY) were coated overnight at 4 °C with 50 µl of 50 µg/ml mono-mannosylated bovine serum albumin (1M-BSA, Dextra Ltd., Reading, UK) in 20 mM sodium bicarbonate buffer (pH 8.5). All subsequent steps were carried out at 25 °C. The 1M-BSA coated culture dishes were blocked and washed with 2 % PBS-BSA. The parallel plate flow chamber apparatus (GlycoTech, Gaithersburg, MD) with gasket size of 2.5 cm x 0.25 cm x 250 µm was fixed on the culture dish and vacuum sealed, followed by equilibration of the system with 0.2 % PBS-BSA with a Harvard Apparatus PHD2000 syringe pump (Instech Laboratories, Inc., Plymouth, PA) and mounted on a Nikon Eclipse Ti phase-contrast microscope (Nikon AG, Zurich, CH). was applied. Bacteria were perfused at wall shear stress between 0.01 pN/µm² and 1.0 pN/µm² and counted after 10 min at 20x magnification (460 µm x 343 µm field of view) at three different points along the centre of the flow path. Exposure time was set according to the shear stress to allow distinction of flowing and bound bacteria. Bacteria were counted automatically with NIS-Elements Advanced Research 3.2 imaging software.

Static bacterial binding and inhibition assay—Nunc MaxiSorp™ 96-well plates (Nunc GmbH, Langenselbold, Germany) were coated with 50 µl of 1M-BSA (50 µg/ml) or ribonuclease B (RNase B, 2 µg/ml, Sigma-Aldrich Co. LLC., Buchs, Switzerland) diluted in 20 mM sodium bicarbonate buffer while shaking overnight at 4 °C. The wells were blocked with 2 % PBS-BSA for 2 h under the same conditions. After washing the blocked wells, the bacterial suspension was added to the plate (50 µl/well). For competitive binding assays, 25 µl/well bacterial suspension was pretreated with a serial dilution of soluble antagonist (25 µl/well) for 15 min before adding it to the plate. The plate was centrifuged for 3 min at 25 °C and 600 x g before incubation for 1 h at 37 °C and 300 rpm. Unbound bacteria were subsequently removed by washing with 0.2 % PBS-BSA and bound bacteria were fixed with 100 µl/well Roti®-Histofix 1 % (Carl Roth GmbH, Karlsruhe, Germany). Bacteria were then labeled for 30 min with the cell-permeable fluorescent dye Syto® 16 (Life Technologies Corp., Zug, Switzerland), diluted to 2 µM in 10 mM HEPES. Fluorescence was measured at 485 nm excitation and 528 nm emission with a BioTek Synergy™ HT fluorometer (BioTek Instruments Inc., Winooski, VT). Static binding assays were performed in duplicate and repeated at least twice. IC₅₀ values from competitive assays were calculated using the Prism

software (GraphPad Software, Inc., La Jolla, CA) by non-linear, four-parameter curve fitting for minimum, maximum, logIC₅₀, and slope.

Transmission electron microscopy—The bacteria from overnight culture were applied on 200-400 mesh per inch copper grids coated with a Formvar[®] support film. The grids were glow discharged shortly before they were treated with the bacterial suspension for 10 minutes. Afterwards, the grids were washed with distilled water and negatively stained with a 0.5 % aqueous uranyl acetate solution for 30 seconds. Following air drying, the samples were examined in a Philips CM 100 transmission electron microscope at an 100 kV acceleration voltage.

Results

Clinical isolates display high variability in the fimH gene—Analysis of the variability in the *fimH* gene from 57 isolates from patients with acute cystitis revealed numerous mutations (Tab. 3), none of which concern the residues in the binding site (Fig. 1). FimH as found in the pathogenic strain J96 (considered as wild type herein) represented the largest population with 14 % of all isolates. The most frequently found mutations are V27A, N70S, S78N. These three mutations together correspond to the FimH of the fecal F18 isolate and were found in 7 % of the isolates. In combination with additional mutations, the V27A, N70S, and S78N were found in 21 % of all tested isolates. The V128M/G (9 %), G66S (9 %), and R166H (10 %) mutations were also found frequently, almost exclusively in combination with the V27A mutation and further, non-consistent mutations. The S62A mutation as found in the highly pathogenic UTI89 clinical isolate was not found in any of the 57 clinical isolates.

TABLE 3

Sequence analysis of 57 UPEC strains isolated from patients with acute cystitis. The residue number, as well as the most frequent mutation are given. FimH^{J96} was considered as wild type.

Residue #	7	27	41	66	68	70	74	78	101
Mutation	N→K	V→A	Q→R	G→S	L→V	N→S	T→I	S→N	K→N
Frequency	1	46	1	5	1	16	3	19	1
Residue #	106	117	118	119	122	128	130	147	163
Mutation	A→V	G→R	A→V	A→V	A→V	V→M	T→A	N→S	V→I
Frequency	2	1	1	3	1	5	1	1	3
Residue #	166	192	202	238	242	273	279		
Mutation	R→H	N→Δ	A→V	V→E	A→V	G→A	Q→K		
Frequency	6	1	1	1	2	2	1		

Although mutations of the *fimH* gene known for altering the binding phenotype were found in this set of clinical isolates, their binding affinity towards mannose did not correlate with mutations (data not shown). This was a consequence of phase variation and the distinct expression of type I pili.

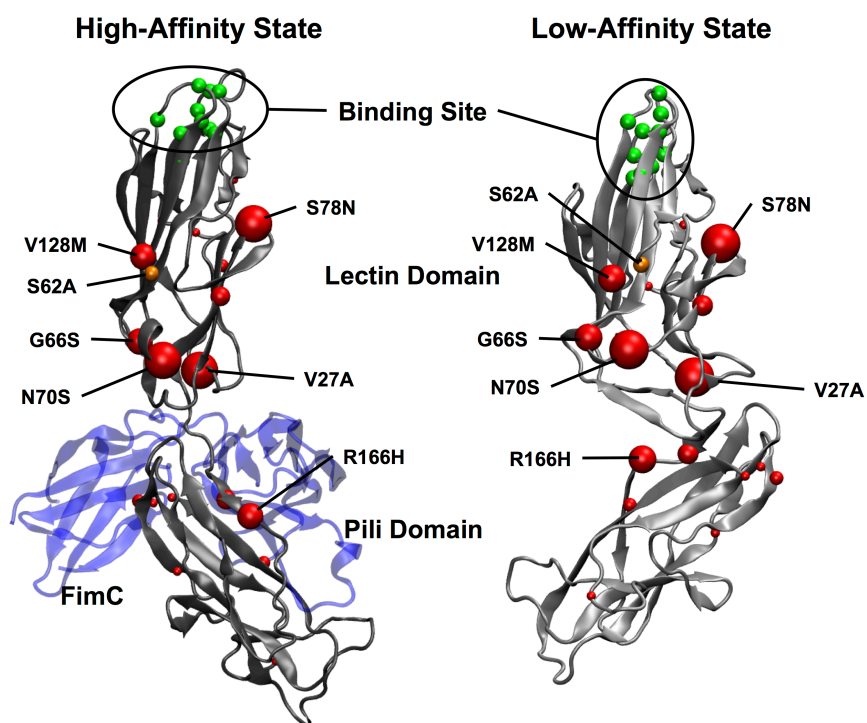


FIGURE 1. **Mutations of the *fimH* gene identified in 57 clinical isolates.** The mutations are mapped as red spheres of the C α atom on the high-affinity conformation (left, PDB code 1QUN) and low-affinity conformation (right, PDB code 3JWN) structures. Sphere size correlates with the frequency of the mutation. The orange sphere corresponds to the S62A mutation found in the clinical isolate UTI89 which is frequently used in *in vitro* and *in vivo* studies. Green spheres represent residues involved in ligand binding (33).

FimH variants found in clinical isolates show differential binding profiles—To overcome the problem of phase-variation, bacteria constitutively expressing type I pili were generated (Fig. 2A). A set of four naturally occurring mutations that cover a wide range from low to strong affinity binding were chosen (Tab. 2). Furthermore, an artificial mutant (A188D, FimH^{high}) that exclusively adopts the high affinity state, was generated as positive control (7,15). To verify the equal expression profile among the isogenic strains, quantitative staining with an anti-type 1 pili antibody was performed, which confirmed the similar expression level of type I pili (Fig. 2B). The mannose binding profile was evaluated in a static assay using 1M-BSA and RNase B, a glycoprotein expressing oligomannose-type structures (Man₅GlcNAc₂ to Man₉GlcNAc₂ N-linked glycans) (Fig. 2C) (34). Bacteria were allowed to bind to glycoprotein coated surfaces and were quantified using a nucleic acid stain and subsequent fluorescence readout. The natural FimH variants exhibited diverse binding to the glycoproteins, while the Δ FimH strain showed no binding towards the mannosylated surface. Binding to RNase B was higher when compared to 1M-BSA due to the presence of α 1-3 and α 1-6 linked oligosaccharides that confer strong binding. While RNase B binding was more uniform, the effect of the mutations was more distinctive for 1M-BSA binding. The artificial

FimH^{high} variant showed the highest binding affinity, followed by the FimH^{UTI89} and FimH^{J96} variants. The FimH^{F18} and FimH^{B126} variants exhibited only weak binding to 1M-BSA.

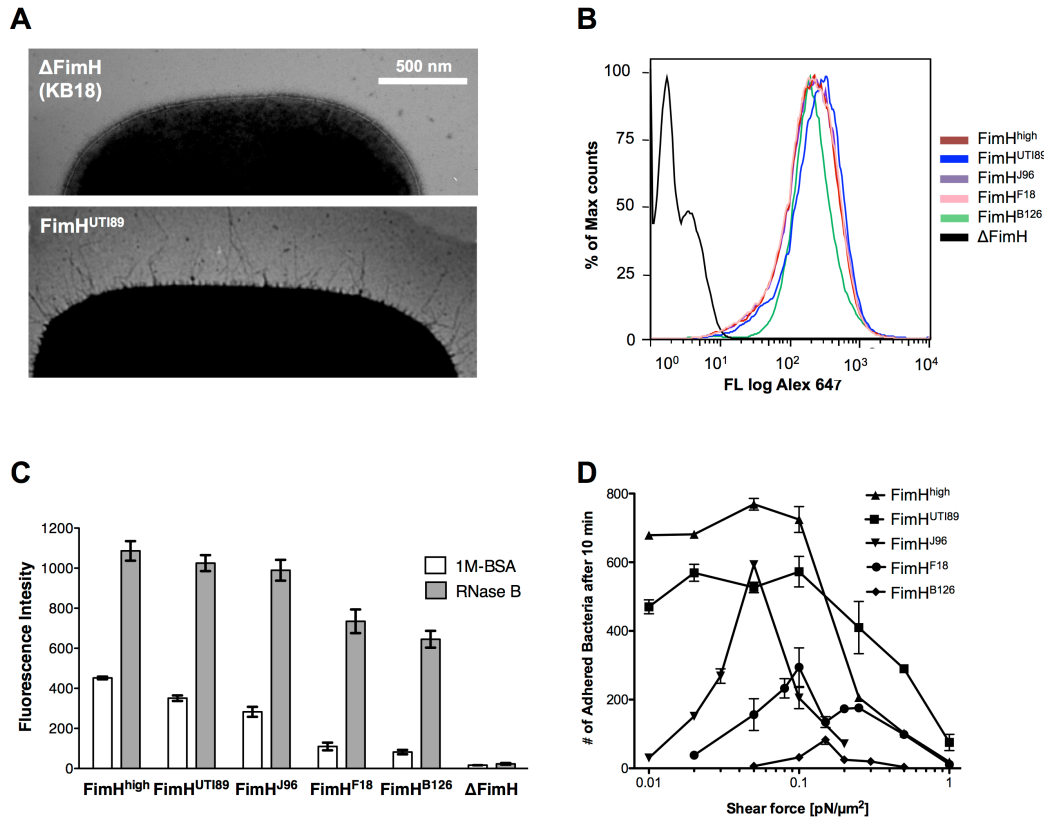


FIGURE 2. Isogenic bacteria expressing naturally occurring FimH variants show distinct binding under static and flow conditions. Electron micrographs (A) show type 1 pili expression of the isogenic strain expressing FimH^{UTI89} (bottom) in comparison to the negative control ΔFimH (top, KB18). Homologous type 1 pili expression was verified by FACS analysis after labeling with antibodies raised against type-1 fimbriae and detection with fluorescently labeled secondary antibody. (B). Static binding (C) was measured on 1M-BSA (50 $\mu\text{g}/\text{ml}$) and RNase B (2 $\mu\text{g}/\text{ml}$) coated surfaces. Bound bacteria were fluorescently labeled with Syto[®]16 (2 μM). The error bars represent the standard error of the mean of two experiments, each performed in duplicate. Binding to 1M-BSA (50 $\mu\text{g}/\text{ml}$) under flow conditions (D) was performed at wall shear stress from 0.01–1.0 pN/ μm^2 . Bound bacteria were counted after 10 minutes of perfusion at three points along the center of the flow path.

We further assessed the binding behavior to 1M-BSA of the FimH variants under physiological flow conditions at wall shear stress of 0.01–1.0 pN/ μm^2 (Fig. 2D). The FimH^{high} variant showed strongest binding that was constant up to a wall shear stress of 0.1 pN/ μm^2 . The FimH^{UTI89} variant exhibited generally decreased attachment and was capable of binding at a wide range of shear force, showing a slight decrease at low shear. The FimH^{J96}, FimH^{F18}, and FimH^{B126} variants displayed catch-bond behavior within the applied range of wall shear force. Consistent with the static assay, FimH^{J96} showed highest attachment followed by FimH^{F18} and FimH^{B126}, the latter only with limited binding. Furthermore, the maximal attachment was shifted towards higher force, with FimH^{J96} peaking at 0.05 pN/ μm^2 ,

FimH^{F18} at 0.1 pN/μm², and FimH^{B126} at 0.15 pN/μm². In conclusion, the selection of FimH variants covers a wide range of binding and is representative for the clinical isolate strains.

Mannoside-based FimH antagonists show differential binding to FimH variants—The potency of FimH antagonists to the clinically relevant FimH variants were evaluated. A set of FimH antagonists representing four different classes of potent α-D-mannopyranosides substituted with various aglycons were chosen (Fig. 3): an alkyl- (*n*-heptyl, **1**) (33), an indolinyphenyl- (**2**), a biphenyl- (**3**) (22,24), and a squaric acid derivative (**4**) substituted mannoside. The analysis was performed with a static cell-based inhibition assay, in which the antagonists compete with a mannosylated surface (Fig. 4A). Since binding to 1M-BSA was too weak for the low affinity FimH variants and therefore unsuitable for proper detection, RNase B coated surfaces were used.

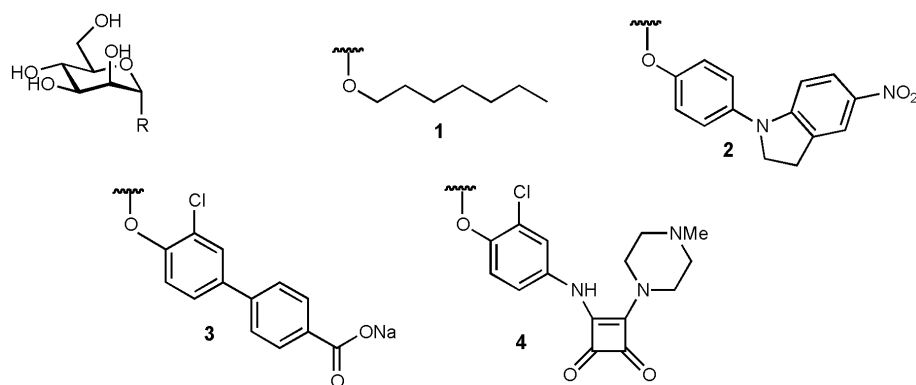


FIGURE 3. **FimH antagonists tested with the FimH variants in the static cell-based competition assay.** α -D-mannopyranoside substituted either with *n*-heptyl (**1**), an indolinyphenyl derivative (**2**), a biphenyl derivative (**3**), and a squaric acid derivative (**4**).

The FimH antagonists **1–4** showed IC_{50} values in the micromolar range for all of the tested FimH variants (Tab. 4). Antagonist **1** was the weakest compound for all types of FimH. Compared to antagonist **1**, antagonists **2** and **3** showed an increase in relative potency (rIP) between ten- to twenty-fold. This increase in affinity was similar for all FimH variants. The squaric acid derivative **4** was the most potent antagonist with an affinity in the low micromolar range. For antagonists **1**, **2**, and **3** we observed a FimH variant-dependent affinity. They exhibited best inhibition towards the locked high affinity FimH^{high} variant, followed by the FimH^{UTI89} and FimH^{I96} variants. The binding affinities towards the low-affinity variants FimH^{F18} and FimH^{B126} were similarly low. The largest difference in IC_{50} value between the FimH variants was observed for antagonist **1**, with a five-fold lower affinity towards FimH^{B126} (97.4 μ M) compared to FimH^{high} (21.8 μ M). Taking only the affinity of antagonist **1** to naturally occurring FimH variants into account, the greatest difference was a two-fold decrease between FimH^{UTI89} (48.7 μ M) and FimH^{B126} (97.4 μ M).

TABLE 4

Inhibition potencies of antagonists to the FimH variants determined with the cell-based inhibition assay. The bacteria ($5 \cdot 10^8$ CFU/mL) were incubated with serial dilutions of the soluble antagonist (1–4) for 15 minutes and allowed to bind to RNase B (2 μ g/ml) coated surfaces under static conditions. All inhibition assays were performed in duplicates and at least twice. The corresponding relative inhibitory potencies (rIP) were calculated for each FimH variant with antagonist 1 as reference compound.

Compound	1		2		3		4	
	IC ₅₀ [μ M]	rIP	IC ₅₀ [μ M]	rIP	IC ₅₀ [μ M]	rIP	IC ₅₀ [μ M]	rIP
FimH ^{high}	21.8 \pm 12.9	1	1.5 \pm 0.5	15	1.5 \pm 0.1	15	1.6 \pm 0.1	14
FimH ^{UTI89}	48.7 \pm 4.9	1	3.0 \pm 0.5	16	3.5 \pm 0.7	14	2.3 \pm 0.2	21
FimH ^{J96}	66.2 \pm 16.2	1	4.4 \pm 0.8	15	3.6 \pm 0.7	18	1.2 \pm 0.1	55
FimH ^{F18}	83.0 \pm 20.0	1	8.4 \pm 0.7	10	4.5 \pm 1.5	18	1.3 \pm 0.1	64
FimH ^{B126}	97.4 \pm 26.2	1	9.9 \pm 1.8	10	5.4 \pm 0.5	18	1.3 \pm 0.6	75

Unlike antagonists 1–3, antagonist 4 exhibited almost uniform binding affinity to all FimH variants, leading to considerably differing rIP values. While antagonist 4 bound with 75-fold higher affinity to FimH^{B126}, only a 21-fold increase was observed for the natural FimH^{UTI89}. Moreover, compound 4 was not more efficacious for the FimH^{high} variant (rIP = 14) than compounds 2 (rIP = 15) and 3 (rIP = 15). The squaric acid derivative 4 was therefore the only FimH antagonist of the herein tested compounds that showed a FimH variant-independent affinity.

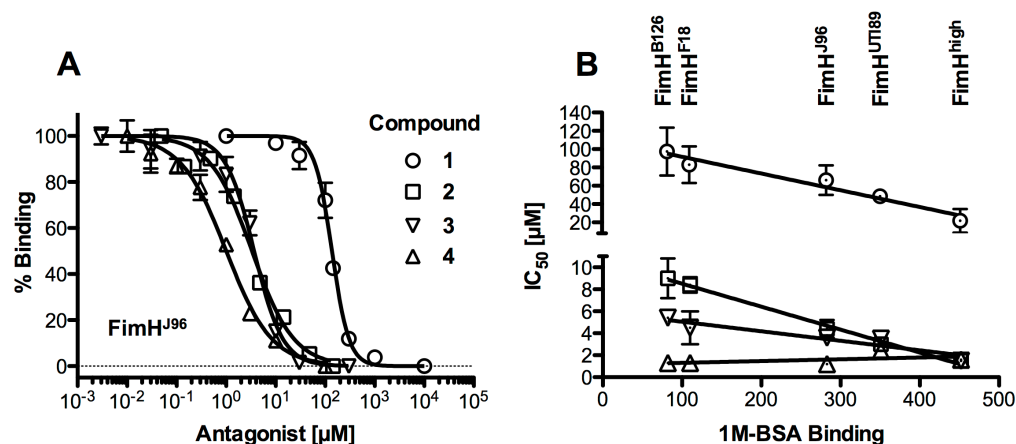


FIGURE 4. FimH antagonists prevent binding of FimH bearing bacteria to a mannosylated surface. Typical IC₅₀ curves as observed in the cell-based inhibition assay on the example of the FimH^{J96} variant (A). Correlation of binding affinity of the FimH variants towards 1M-BSA with the IC₅₀ values obtained for the four compound classes (B). Higher 1M-BSA binding correlated with increase affinity towards antagonists 1–3. This correlation was not observed for antagonist 4, for which similar IC₅₀ values were observed for all FimH variants.

Discussion

The colonization of the urinary tract by UPECs that ultimately leads to acute cystitis is a serious health risk. Since bacterial resistance due to overuse of antibiotics is increasingly problematic (35), drug discovery is focused on alternative treatment possibilities for UTIs. The mannose-specific bacterial adhesin FimH, being essential for UPEC pathogenicity (36-38), is a suitable therapeutic target. Recent efforts have led to α -D-mannopyranosides substituted with various aglycons that show micromolar to nanomolar affinity, depending on the employed assay format (22,24,33,39-41). Such anti-adhesive therapy is believed not to trigger bacterial resistances, since they act neither bacteriostatic, nor bacteriolytic. Numerous assay formats exist with which FimH antagonists are evaluated. For cell-free assays (target-based competitive binding assays (25), isothermal titration calorimetry (24,42), and surface plasmon resonance (33), recombinantly expressed and purified FimH_{LD} is employed. However, due to the missing of the FimH_{PD}, antagonists only encounter the high-affinity state of FimH, which does not correctly reflect the situation *in vivo*. Unfortunately, expression of full length FimH, *i.e.* LD and PD, is not possible due to the missing donor strand rendering FimH alone unstable. Constructs bearing FimH_{LD+PD} in conjunction with FimC are stable, however, in such dimers the high-affinity state of FimH is adopted (37). FimH antagonists have also extensively been evaluated using cell-based *in vitro* assays, *e.g.* hemagglutinin assays (27,39,43,44) and a flow cytometry based assay (45). In these assays, the highly pathogenic clinical isolate UTI89 is employed due to its high degree of binding and ease of handling. A comparison of target- and cell-based assays used for testing antagonists **1–4** is shown in Tab. 5. Not only do the absolute numbers (nM vs μ M) differ significantly between the target- and cell-based assays, also the relative potencies differ considerably. With the exception of compound **2** in the aggregometry assay, the three cell-based assays correlate well. For example, compound **4** shows merely a three-fold improvement over compound **1** in the target-based assay, while it is 14 to 35-fold more potent in cell-based assays. This might be a consequence of the proper display of full length FimH with its ability of adopting the high- and the low-affinity conformation.

TABLE 5

Comparison of relative inhibitory potencies (rIP) for compounds 2–4 in relation to compound 1 determined in a target-based competitive binding assay, a flow cytometry assay, an erythrocyte aggregometry assay, and the cell-based inhibition assays presented herein. The range of rIP values for the various FimH variants are given in the cell-based inhibition assay. The affinity for compound 1 is given for each assay.

	Target	Competitor	1 rIP	2 rIP	3 rIP	4 rIP
Target-based assay (22,24,25,40)	FimH _{LD} Recomb.	Trimannose- polymer	1 (55 nM)	3.7	11.1	3.1
Flow cytometry assay (22,24,40,45)	UTI89 Clin. Isol.	Bladder epithelial cells (5637)	1 (4 µM)	27.8	11.8	35.7
Aggregometry assay (22,27,40)	UTI89 Clin. Isol.	Guinea Pig Erythrocytes	1 (77 µM)	2.9	14.6	15.4
Cell-based inhibition assay (This study)	Isogenic FimH var.	RNase B (Man ₅ -Man ₉)	1 (22–78 µM)	10–16	14–18	14–75

The UTI89 clinical isolate is not only used in cell-based *in vitro* assays, it is also the strain of choice for *in vivo* efficacy studies (9,22,40). The S62A mutation in the *fimH* gene of UTI89 is crucial for its high binding affinity (30), while otherwise being identical to the low-affinity FimH^{F18} variant. Indeed, FimH^{UTI89} displayed flow adhesion properties distinct from FimH^{J96}, FimH^{F18}, and FimH^{B126}, *i.e.* no catch-bond behavior within the applied range (Fig. 2D). The sequence analysis of 57 clinical UPEC strains revealed that not a single strain possessed the S62A mutation. This is also the case in previously published studies (11,46). While most isolates possess the FimH^{J96} variant, many have the V27A mutation, which is known to stabilize the low affinity conformation and weaken mannose binding (47). Other mutations, *e.g.* the M128V mutation, even further weaken mannose binding. Nonetheless, strains bearing low-affinity FimH appear to be able of causing UTIs since they benefit from reduced affinity to the Tamm-Horsfall protein. It was therefore of great importance to evaluate the efficacy of FimH antagonists towards a set of naturally occurring FimH variants. The FimH antagonists belonging to four separate classes were all capable of binding to all of the FimH variants with distinct mannose-binding phenotypes and are able of inhibiting their adhesion to a mannosylated surface. The affinity of compounds 1–3 thereby correlated with the binding phenotype of the FimH variants. These antagonists showed highest affinity towards the high 1M-BSA binding FimH variant (FimH^{high}) and a gradually decreased binding to the low 1M-BSA binding variants (Fig. 4B). This was however not true for the squaric acid derivative 4, which showed similar affinity towards all FimH variants resulting in great differences in rIP values compared to compound 1. Since the possibility of covalent binding for squaric acid derivatives to FimH has been excluded (48), perhaps an alternative

mechanism of action of **4**, *e.g.* an alternative binding mode that does not discriminate between the two affinity states of FimH might explain this behavior.

In summary, this study systematically demonstrates the efficacy of FimH antagonists towards several naturally occurring FimH variants for the first time. While there is a correlation between the affinity of antagonists and the binding phenotype of the various FimH types, these differences are small. We therefore postulate that potent FimH antagonists are suitable for treatment and prevention of UTI in which a wide spectrum of FimH variants are involved.

Literature

1. Ronald, A. (2002) *Am. J. Med.* **113**, 14S-19S
2. Mulvey, M. A. (2002) *Cell. Microbiol.* **4**, 257-271
3. Xie, B., Zhou, G., Chan, S. Y., Shapiro, E., Kong, X. P., Wu, X. R., Sun, T. T., and Costello, C. E. (2006) *J. Biol. Chem.* **281**, 14644-14653
4. Zhou, G., Mo, W. J., Sebbel, P., Min, G. W., Neubert, T. A., Glockshuber, R., Wu, X. R., Sun, T. T., and Kong, X. P. (2001) *J. Cell Sci.* **114**, 4095-4103
5. Choudhury, D., Thompson, A., Stojanoff, V., Langermann, S., Pinkner, J., Hultgren, S. J., and Knight, S. D. (1999) *Science* **285**, 1061-1066
6. Le Trong, I., Aprikian, P., Kidd, B. A., Forero-Shelton, M., Tchesnokova, V., Rajagopal, P., Rodriguez, V., Interlandi, G., Klevit, R., Vogel, V., Stenkamp, R. E., Sokurenko, E. V., and Thomas, W. E. (2010) *Cell* **141**, 645-655
7. Aprikian, P., Tchesnokova, V., Kidd, B., Yakovenko, O., Yarov-Yarovoy, V., Trinchina, E., Vogel, V., Thomas, W., and Sokurenko, E. (2007) *J. Biol. Chem.* **282**, 23437-23446
8. Yakovenko, O., Tchesnokova, V., Aprikian, P., Forero, M., Vogel, V., Sokurenko, E., and Thomas, W. (2007) *Biophys. J.*, 348A-349A
9. Wellens, A., Garofalo, C., Nguyen, H., Van Gerven, N., Slattegard, R., Hernalsteens, J. P., Wyns, L., Oscarson, S., De Greve, H., Hultgren, S., and Bouckaert, J. (2008) *PLoS ONE* **3**
10. Tchesnokova, V., Aprikian, P., Yakovenko, O., Larock, C., Kidd, B., Vogel, V., Thomas, W., and Sokurenko, E. (2008) *J. Biol. Chem.* **283**, 7823-7833
11. Sokurenko, E. V., Chesnokova, V., Dykhuizen, D. E., Ofek, I., Wu, X. R., Krogfelt, K. A., Struve, C., Schembri, M. A., and Hasty, D. L. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95**, 8922-8926
12. Sokurenko, E. V., Feldgarden, M., Trintchina, E., Weissman, S. J., Avagyan, S., Chattopadhyay, S., Johnson, J. R., and Dykhuizen, D. E. (2004) *Mol. Biol. Evol.* **21**, 1373-1383
13. Weissman, S. J., Beskhlebnaya, V., Chesnokova, V., Chattopadhyay, S., Stamm, W. E., Hooton, T. M., and Sokurenko, E. V. (2007) *Infect. Immun.* **75**, 3548-3555
14. Sokurenko, E. V., Courtney, H. S., Maslow, J., Siitonen, A., and Hasty, D. L. (1995) *J. Bacteriol.* **177**, 3680-3686
15. Anderson, B. N., Ding, A. M., Nilsson, L. M., Kusuma, K., Tchesnokova, V., Vogel, V., Sokurenko, E. V., and Thomas, W. E. (2007) *J. Bacteriol.* **189**, 1794-1802
16. Nilsson, L. M., Thomas, W. E., Sokurenko, E. V., and Vogel, V. (2006) *Appl. Environ. Microbiol.* **72**, 3005-3010
17. Ofek, I., Mirelman, D., and Sharon, N. (1977) *Nature* **265**, 623-625
18. Sharon, N. (2006) *Bba-Gen Subjects* **1760**, 527-537
19. Ernst, B., and Magnani, J. L. (2009) *Nat. Rev. Drug Discov.* **8**, 661-677
20. Eden, C. S., Freter, R., Hagberg, L., Hull, R., Hull, S., Leffler, H., and Schoolnik, G. (1982) *Nature* **298**, 560-562
21. Aronson, M., Medalia, O., Schori, L., Mirelman, D., Sharon, N., and Ofek, I. (1979) *J. Infect. Dis.* **139**, 329-332
22. Klein, T., Abgottspon, D., Wittwer, M., Rabbani, S., Herold, J., Jiang, X. H., Kleeb, S., Luthi, C., Scharenberg, M., Bezencon, J., Gubler, E., Pang, L. J., Smiesko, M., Cutting, B., Schwardt, O., and Ernst, B. (2010) *J. Med. Chem.* **53**, 8627-8641
23. Sperling, O., Dubber, M., and Lindhorst, T. K. (2007) *Carbohydr. Res.* **342**, 696-703
24. Pang, L., Kleeb, S., Lemme, K., Rabbani, S., Scharenberg, M., Zalewski, A., Schadler, F., Schwardt, O., and Ernst, B. (2012) *Chemmedchem* **7**, 1404-1422

25. Rabbani, S., Jiang, X. H., Schwardt, O., and Ernst, B. (2010) *Anal. Biochem.* **407**, 188-195
26. Bouckaert, J., Mackenzie, J., de Paz, J. L., Chipwaza, B., Choudhury, D., Zavialov, A., Mannerstedt, K., Anderson, J., Pierard, D., Wyns, L., Seeberger, P. H., Oscarson, S., De Greve, H., and Knight, S. D. (2006) *Mol. Microbiol.* **61**, 1556-1568
27. Abgottspon, D., Rolli, G., Hosch, L., Steinhuber, A., Jiang, X. H., Schwardt, O., Cutting, B., Smiesko, M., Jenal, U., Ernst, B., and Trampuz, A. (2010) *J. Microbiol. Methods* **82**, 249-255
28. Hultgren, S. J., Schwan, W. R., Schaeffer, A. J., and Duncan, J. L. (1986) *Infect. Immun.* **54**, 613-620
29. Sokurenko, E. V., Courtney, H. S., Ohman, D. E., Klemm, P., and Hasty, D. L. (1994) *J. Bacteriol.* **176**, 748-755
30. Stahlhut, S. G., Tchesnokova, V., Struve, C., Weissman, S. J., Chattopadhyay, S., Yakovenko, O., Aprikian, P., Sokurenko, E. V., and Krogfelt, K. A. (2009) *J. Bacteriol.* **191**, 6592-6601
31. Thomas, W. E., Trintchina, E., Forero, M., Sokurenko, E. V., and Vogel, V. (2002) *Mol. Biol. Cell* **13**, 404A-404A
32. Thomas, W. E., Nilsson, L. M., Forero, M., Sokurenko, E. V., and Vogel, V. (2004) *Mol. Microbiol.* **53**, 1545-1557
33. Bouckaert, J., Berglund, J., Schembri, M., De Genst, E., Cools, L., Wuhrer, M., Hung, C. S., Pinkner, J., Slattegard, R., Zavialov, A., Choudhury, D., Langermann, S., Hultgren, S. J., Wyns, L., Klemm, P., Oscarson, S., Knight, S. D., and De Greve, H. (2005) *Mol. Microbiol.* **55**, 441-455
34. van Rooijen, J. J. M., Voskamp, A. F., Kamerling, J. P., and Vliegenthart, J. F. G. (1999) *Glycobiology* **9**, 21-30
35. Sanchez, G. V., Master, R. N., Karlowsky, J. A., and Bordon, J. M. (2012) *Antimicrob. Agents Chemother.* **56**, 2181-2183
36. Connell, H., Agace, W., Klemm, P., Schembri, M., Marild, S., and Svanborg, C. (1996) *Proc. Natl. Acad. Sci. U. S. A.* **93**, 9827-9832
37. Hung, C. S., Bouckaert, J., Hung, D., Pinkner, J., Widberg, C., DeFusco, A., Auguste, C. G., Strouse, R., Langermann, S., Waksman, G., and Hultgren, S. J. (2002) *Mol. Microbiol.* **44**, 903-915
38. Wright, K. J., Seed, P. C., and Hultgren, S. J. (2007) *Cell. Microbiol.* **9**, 2230-2241
39. Han, Z. F., Pinkner, J. S., Ford, B., Obermann, R., Nolan, W., Wildman, S. A., Hobbs, D., Ellenberger, T., Cusumano, C. K., Hultgren, S. J., and Janetka, J. W. (2010) *J. Med. Chem.* **53**, 4779-4792
40. Jiang, X. H., Abgottspon, D., Kleeb, S., Rabbani, S., Scharenberg, M., Wittwer, M., Haug, M., Schwardt, O., and Ernst, B. (2012) *J. Med. Chem.* **55**, 4700-4713
41. Schwardt, O., Rabbani, S., Hartmann, M., Abgottspon, D., Wittwer, M., Kleeb, S., Zalewski, A., Smiesko, M., Cutting, B., and Ernst, B. (2011) *Bioorg. Med. Chem.* **19**, 6454-6473
42. Wellens, A., Lahmann, M., Touaibia, M., Vaucher, J., Oscarson, S., Roy, R., Remaut, H., and Bouckaert, J. (2012) *Biochemistry (Mosc.)* **51**, 4790-4799
43. Cusumano, C. K., Pinkner, J. S., Han, Z. F., Greene, S. E., Ford, B. A., Crowley, J. R., Henderson, J. P., Janetka, J. W., and Hultgren, S. J. (2011) *Sci. Transl. Med.* **3**
44. Han, Z. F., Pinkner, J. S., Ford, B., Chorell, E., Crowley, J. M., Cusumano, C. K., Campbell, S., Henderson, J. P., Hultgren, S. J., and Janetka, J. W. (2012) *J. Med. Chem.* **55**, 3945-3959
45. Scharenberg, M., Abgottspon, D., Cicek, E., Jiang, X. H., Schwardt, O., Rabbani, S., and Ernst, B. (2011) *Assay Drug Dev. Technol.* **9**, 455-464

46. Hommais, F., Gouriou, S., Amarin, C., Bui, H., Rahimy, M. C., Picard, B., and Denamur, E. (2003) *Infect. Immun.* **71**, 3619-3622
47. Thomas, W. E., Trintchina, E., Forero, M., Vogel, V., and Sokurenko, E. V. (2002) *Cell* **109**, 913-923
48. Grabosch, C., Hartmann, M., Schmidt-Lassen, J., and Lindhorst, T. K. (2011) *Chembiochem* **12**, 1066-1074

Supplementary Material

SUPPLEMENTAL TABLE S1. List of clinical UPEC isolates with *fimH* gene sequence and antibiotic resistance analysis. FimH^{J96} was considered as wild type. Amx: Amoxicillin, Clv: Clavulanic acid, Ctr: Cotrimoxazole, Cip: Ciprofloxacin. N.d. = not determined.

Name	Mutations	Resistances
B1 1	N7K / V27A / V128M	n.d.
B1 2	V27A	None
B1 3	V27A	None
B1 4	V27A	n.d.
B1 5	V27A / S78N / A242V	Amx/Clv
B1 7	G117R	n.d.
B1 10	No mutations	n.d.
B1 11	V27A	None
B1 12	No mutations	Ctr
B1 14	No mutations	Amx/Clv
B1 15	V27A / N70S	None
B1 16	V27A / R166H	Amx/Clv, Ctr, Cip
B1 17	V27A / N70S / S78N	None
B1 18	No mutations	None
B1 19	V27A / G66S / A202V	None
B1 20	No mutations	None
B1 23	V27A / V128G / R166H	None
B1 24	V27A / Q41R / V128M	None
B1 26	V27A / V128M	None
B1 27	N70S	None
B1 29	V27A / V163I / T130A	None
B1 31	G66S / S78N / A242V	Ctr
B1 32	V27A / S78N	None
B1 33	V27A / G66S / N70S / S78N	Cip
B1 34	V27A / N70S / S78N	None
B1 35	No mutations	None
B1 37	V27A / S78N	None
B1 38	V27A / A118V	Cip
B1 39	V27A / R166H	Cip
B1 40	V27A / V128M	None
B1 45	V27A / L68V / N70S / S78N	None
B1 46	V27A / N70S / S78N / V163A	Ctr
B1 47	V27A / N70S / T74I / S78N / V238E	Amx/Clv
B1 49	V27A / A119V / G273A	Amx/Clv, Ctr, Cip

B1 51	V27A	Amx/Clv, Cip
B1 53	V27A / N70S / S78N	Amx/Clv
B1 55	V27A / N70S / S78N / N147S	None
B1 56	V27A / N70S / T74I / S78N / V163A	Ctr
B1 57	V27A	None
B1 58	V27A / N192Δ	Ctr
B1 59	No mutations	None
B1 60	V27A / N70S / S78N / A106V / A122V	None
B1 61	V27A / S78N	None
B3 1	V27A / R166H	n.d.
B3 2	V27A / Q279K	None
B3 3	V27A / N70S / S78N / K101N	None
B3 4	V27A / T74I	None
B3 5	V27A / N70S / S78N / A106V	None
B3 6	V27A / N70S / S78N	None
B3 7	V27A / A119V	None
B3 8	V27A / G66S / N70S / S78N	Ctr
B3 13	V27A / R166H	None
B3 14	V27A / G66S / N70S / S78N	None
B3 15	V27A / A119V / G273A	None
B3 16	V27A	Cip
B3 17	No mutations	None
B3 18	V27A / R166H	None